https://www.facebook.com/GrinchaZ

Chapter 1

Sets and Probability

THE CONCEPT OF A SET

At the foundations of probability and statistics, and of mathematics in general, is the concept of a *set*. A set can be thought of as a collection of objects, called *members* or *elements* of the set. In general, unless otherwise specified, we denote a set by a capital letter such as A, B, C, and an element by a lower case letter such as a, b. Synonyms for set are *class, aggregate* and *collection.*

If an element a belongs to a set C we write $a \in C$. If a does not belong to C we write $a \notin C$. If both a and b belong to C we write $a, b \in C$. In order for a set to be *well-defined,* as we shall always assume, we must be able to determine whether a particular object does or does not belong to the set.

A set can be defined by actually listing its elements or, if this is not possible, by describing some property held by all members and by no nonmembers. The first is called the *roster method* and the second is called the *property method.*

EXAMPLE 1.1. The set of all vowels in the English alphabet can be defined by the roster method as $\{a, e, i, o, u\}$ or by the property method as $\{x \mid x \text{ is a vowel}\}$, read "the set of all elements x such that x is a vowel" where the vertical line $\mid$ is read "such that" or "given that."

EXAMPLE 1.2. The set $\{x \mid x \text{ is a triangle in a plane}\}$ is the set of all triangles in a plane. Note that the roster method cannot be used here.

EXAMPLE 1.3. If we toss a pair of ordinary dice the possible "numbers" or "points" which can come up on the uppermost face of each die are elements of the set $\{1, 2, 3, 4, 5, 6\}$.

SUBSETS

If each element of a set A also belongs to a set B we call A a *subset* of B, written $A \subset B$ or $B \supset A$ and read "A is contained in B" or "B contains A" respectively. It follows that for all sets A we have $A \subset A$.

If $A \subset B$ and $B \subset A$ we call A and B *equal* and write $A = B$. In such case A and B have exactly the same elements.

If A is not equal to B, i.e. if A and B do not have exactly the same elements, we write $A \neq B$.

If $A \subset B$ but $A \neq B$ we call A a *proper subset* of B.

EXAMPLE 1.4. $\{a, i, u\}$ is a proper subset of $\{a, e, i, o, u\}$.

EXAMPLE 1.5. $\{i, o, a, u, e\}$ is a subset, but not a proper subset, of $\{a, e, i, o, u\}$, because the two sets are equal. Note that a mere rearrangement of elements does not change the set.

EXAMPLE 1.6. In tossing a die the possible outcomes where the die comes up "even" are elements of the set {2, 4, 6}, which is a (proper) subset of the set of all possible outcomes {1, 2, 3, 4, 5, 6}.

The following theorem is true for any sets A, B, C.

Theorem 1-1: If $A \subset B$ and $B \subset C$, then $A \subset C$.

UNIVERSAL SET AND EMPTY SET

For many purposes we restrict our discussion to subsets of some particular set called the *universe of discourse*, or briefly, *universe*. It is also referred to as the *universal set* or *space* and denoted by $\mathcal{U}$. The elements of a space are often called *points* of the space.

It is useful to consider a set having no elements at all. This is called the *empty set* or *null set* and is denoted by $\emptyset$. It is a subset of any set.

EXAMPLE 1.7. An important set already familiar is the set $\mathcal{R}$ of *real numbers* such as 3, -2, $\sqrt{2}$, π, which can be represented by points on a *real line* such as the x-axis. If a and b are real numbers such that $a < b$, the subsets $\{x \mid a \leqq x \leqq b\}$ and $\{x \mid a < x < b\}$ of $\mathcal{R}$ (often denoted briefly by $a \leqq x \leqq b$ and $a < x < b$) are called *closed* and *open intervals* respectively. Subsets such as $\{x \mid a \leqq x < b\}$ or $\{x \mid a < x \leqq b\}$ are then called *half open* or *half closed intervals*.

EXAMPLE 1.8. The set of all real numbers x such that $x^2 = -1$, written $\{x \mid x^2 = -1\}$, is the null or empty set since there are no real numbers whose squares are equal to -1. If we include complex numbers, however, the set is not empty.

EXAMPLE 1.9. If we toss a die the set of all possible outcomes is the universe {1, 2, 3, 4, 5, 6}. The set of outcomes consisting of faces 7 or 11 on a single die is the null set.

VENN DIAGRAMS

A universe $\mathcal{U}$ can be represented geometrically by the set of points inside a rectangle. In such case subsets of $\mathcal{U}$ (such as A and B shown shaded in Fig. 1-1) are represented by sets of points inside circles. Such diagrams, called *Venn diagrams*, often serve to provide geometric intuition regarding possible relationships between sets.

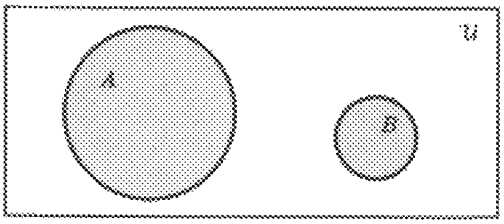


Fig. 1-1

SET OPERATIONS

1. **Union.** The set of all elements (or points) which belong to either A or B or both A and B is called the *union* of A and B and is denoted by $A \cup B$ (shaded in Fig. 1-2).

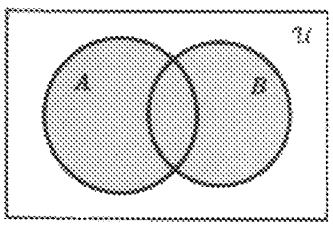


Fig. 1-2

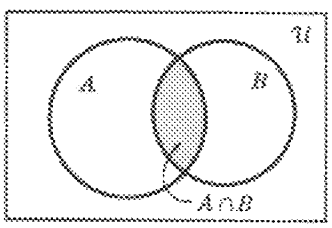


Fig. 1-3

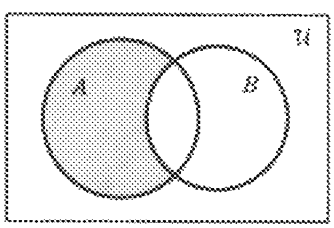


Fig. 1-4

2. **Intersection.** The set of all elements which belong to both A and B is called the *intersection* of A and B and is denoted by $A \cap B$ (shaded in Fig. 1-3).

   Two sets A and B such that $A \cap B = \emptyset$, i.e. which have no elements in common, are called *disjoint sets*. In Fig. 1-1, A and B are disjoint.

3. **Difference.** The set consisting of all elements of A which do not belong to B is called the *difference* of A and B, denoted by $A - B$ (shaded in Fig. 1-4).

4. **Complement.** If $B \subset A$ then $A - B$ is called the *complement of B relative to A* and is denoted by B'_A (shaded in Fig. 1-5). If $A = \mathcal{U}$, the universal set, we refer to $\mathcal{U} - B$ as simply the *complement of B* and denote it by B' (shaded in Fig. 1-6). The complement of $A \cup B$ is denoted by $(A \cup B)'$.

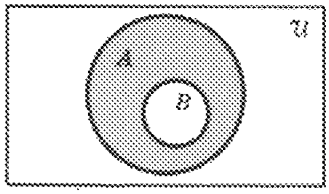


Fig. 1-5

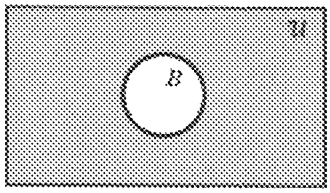


Fig. 1-6

SOME THEOREMS INVOLVING SETS

Theorem 1-2:	$A \cup B = B \cup A$	Commutative law for unions
Theorem 1-3:	$A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$	Associative law for unions
Theorem 1-4:	$A \cap B = B \cap A$	Commutative law for intersections
Theorem 1-5:	$A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$	Associative law for intersections
Theorem 1-6:	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	First distributive law
Theorem 1-7:	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Second distributive law
Theorem 1-8:	$A - B = A \cap B'$	
Theorem 1-9:	If $A \subset B$, then $A' \supset B'$ or $B' \subset A'$	
Theorem 1-10:	$A \cup \emptyset = A, \quad A \cap \emptyset = \emptyset$	
Theorem 1-11:	$A \cup \mathcal{U} = \mathcal{U}, \quad A \cap \mathcal{U} = A$	
Theorem 1-12a:	$(A \cup B)' = A' \cap B'$	De Morgan's first law
Theorem 1-12b:	$(A \cap B)' = A' \cup B'$	De Morgan's second law
Theorem 1-13:	$A = (A \cap B) \cup (A \cap B')$	for any sets A and B

Theorems 1-12*a*, 1-12*b* and 1-13 can be generalized (see Problems 1.69 and 1.74).

PRINCIPLE OF DUALITY

Any true result involving sets is also true if we replace unions by intersections, intersections by unions, sets by their complements and if we reverse the *inclusion symbols* $\subset$ and $\supset$.

RANDOM EXPERIMENTS

We are all familiar with the importance of experiments in science and engineering. A fundamental principle involved is that if we perform such experiments repeatedly under very nearly identical conditions we arrive at results which are essentially the same.

There are, however, experiments in which results will not be essentially the same even though conditions may be nearly identical. Such experiments are called *random experiments.* The following are some examples.

EXAMPLE 1.10. If we toss a coin the result of the experiment is that it will either come up "tails," symbolized by T (or 0), or "heads," symbolized by H (or 1), i.e. one of the elements of the set $\{H, T\}$ (or $\{0, 1\}$).

EXAMPLE 1.11. If we toss a die the result of the experiment is that it will come up with one of the numbers in the set $\{1, 2, 3, 4, 5, 6\}$.

EXAMPLE 1.12. If we toss a coin twice, the results can be indicated by $\{HH, HT, TH, TT\}$, i.e. both heads, head on first and tail on second, etc.

EXAMPLE 1.13. If we have a machine which makes bolts, the result of the experiment is that some may be defective. Thus when a bolt is made it will be a member of the set {defective, nondefective}.

EXAMPLE 1.14. If an experiment consists in measuring "lifetimes" of electric light bulbs produced by a company, then the result of the experiment is a time t in hours which lies in some interval—say, $0 \leq t \leq 4000$—where we assume that no bulb lasts more than 4000 hours.

SAMPLE SPACES

A set $\mathcal{S}$ which consists of all possible outcomes of a random experiment is called a *sample space* and each outcome is called a *sample point.* Often there will be more than one sample space which can describe outcomes of an experiment but there is usually only one which will provide the most information. Note that $\mathcal{S}$ corresponds to the universal set.

EXAMPLE 1.15. If we toss a die, one sample space or set of all possible outcomes is given by $\{1, 2, 3, 4, 5, 6\}$ while another is {odd, even}. It is clear, however, that the latter would not be adequate to determine, for example, whether an outcome is divisible by 3.

It is often useful to portray a sample space graphically. In such case it is desirable to use numbers in place of letters whenever possible.

EXAMPLE 1.16. If we toss a coin twice and use 0 to represent tails and 1 to represent heads the sample space (see Example 1.12) can be portrayed by the points in Fig. 1-7 where, for example, (0, 1) represents tails on first toss and heads on second toss, i.e. TH.

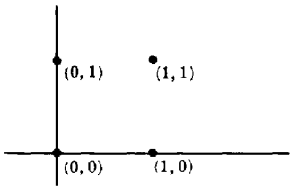


Fig. 1-7

If a sample space has a finite number of points, as in Example 1.16, it is called a *finite sample space.* If it has as many points as there are natural numbers $1, 2, 3, \ldots$, it is called a *countably infinite sample space.* If it has as many points as there are in some interval on the x-axis, such as $0 \leq x \leq 1$, it is called a *noncountably infinite sample space.* A sample space which is finite or countably infinite is often called a *discrete sample space,* while one which is noncountably infinite is called a *nondiscrete* or *continuous sample space.*

EVENTS

An *event* is a subset A of the sample space $\mathcal{S}$, i.e. it is a set of possible outcomes. If the outcome of an experiment is an element of A we say that the event A *has occurred.* An event consisting of a single point of $\mathcal{S}$ is often called a *simple* or *elementary event.*

EXAMPLE 1.17. If we toss a coin twice, the event that only one head comes up is the subset of the sample space which consists of points (0, 1) and (1,0), as indicated in Fig. 1-8.

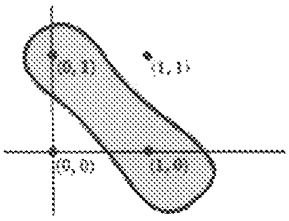


Fig. 1-8

As particular events we have $\mathscr{S}$ itself, which is the *sure* or *certain event* since an element of $\mathscr{S}$ must occur, and the empty set $\emptyset$, which is called the *impossible event* because an element of $\emptyset$ cannot occur.

Since events are sets it is clear that statements concerning events can be translated into the language of set theory and conversely. In particular we have an *algebra of events* corresponding to the algebra of sets indicated on page 3.

By using set operations on events in $\mathscr{S}$ we can obtain other events in $\mathscr{S}$. Thus if A and B are events, then

1. $A \cup B$ is the event "either A or B or both."
2. $A \cap B$ is the event "both A and B."
3. A' is the event "not A."
4. $A - B$ is the event "A but not B."

If the sets corresponding to events A and B are disjoint, i.e. $A \cap B = \emptyset$, we often say that the events are *mutually exclusive*. This means that they cannot both occur.

EXAMPLE 1.18. Referring to the experiment of tossing a coin twice let A be the event "at least one head occurs" and B the event "the second toss results in a tail." Then $A = \{HT, TH, HH\}$, $B = \{HT, TT\}$ and so we have

$$A \cup B = \{HT, TH, HH, TT\} = \mathscr{S} \qquad A \cap B = \{HT\}$$

$$A' = \{TT\} \qquad A - B = \{TH, HH\}$$

THE CONCEPT OF PROBABILITY

In any random experiment there is always uncertainty as to whether a particular event will or will not occur. As a measure of the *chance* or *probability* with which we can expect the event to occur it is convenient to assign a number between 0 and 1. If we are sure or certain that the event will occur we say that its probability is 100% or 1, but if we are sure that the event will not occur we say that its probability is zero. If for example the probability is $\frac{1}{4}$, we would say that there is a 25% chance it will occur and a 75% chance that it will not occur. Equivalently we can say that the *odds* against its occurrence are 75% to 25% or 3 to 1.

There are two important procedures by means of which we can get estimates for the probability of an event.

1. **Classical or a priori approach.** If an event can occur in h different ways out of a total number of n possible ways, all of which are equally likely, then the probability of the event is h/n.

EXAMPLE 1.19. Suppose we want the probability that a head turns up in a single toss of a coin. Since there are two equally likely ways in which the coin can come up, namely "heads" and "tails" (assuming it does not roll away or stand on its edge), and of these two ways a head can arise in only one way, we reason that the required probability is 1/2. In arriving at this we assume that the coin is *fair*, i.e. not *loaded* in any way.

Ziyad Asaad

2. **Frequency or a posteriori approach.** If after n repetitions of an experiment, where n is very large, an event is observed to occur in h of these, then the probability of the event is h/n. This is also called the *empirical probability* of the event.

EXAMPLE 1.20. If we toss a coin 1000 times and find that it comes up heads 532 times we estimate the probability of a head to be $532/1000 = 0.532$.

Both the classical and frequency approaches have serious difficulties, the first because the words "equally likely" are vague and the second because of the vagueness of the "large number" involved. Because of these difficulties mathematicians have in recent years been led to an *axiomatic approach* to probability using sets.

THE AXIOMS OF PROBABILITY

Suppose we have a sample space $\mathcal{S}$. If $\mathcal{S}$ is discrete all subsets correspond to events and conversely, but if $\mathcal{S}$ is continuous only special subsets (called *measurable*) correspond to events. To each event A in the class $\mathcal{C}$ of events we associate a real number $P(A)$, i.e. P is a real-valued function defined on $\mathcal{C}$. Then P is called a *probability function*, and $P(A)$ the *probability* of the event A, if the following axioms are satisfied:

Axiom 1. For every event A in the class $\mathcal{C}$

$$P(A) \geqq 0 \tag{1}$$

Axiom 2. For the sure or certain event $\mathcal{S}$ in the class $\mathcal{C}$

$$P(\mathcal{S}) = 1 \tag{2}$$

Axiom 3. For any number of mutually exclusive events $A_1, A_2, \ldots$ in the class $\mathcal{C}$

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots \tag{3}$$

In particular, for only two mutually exclusive events A_1, A_2,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \tag{4}$$

SOME IMPORTANT THEOREMS ON PROBABILITY

From the above axioms we can now prove various theorems on probability which are important in further work.

***Theorem 1-14*:** If $A_1 \subset A_2$ then $P(A_1) \leqq P(A_2)$ and $P(A_2 - A_1) = P(A_2) - P(A_1)$.

***Theorem 1-15*:** For every event A

$$0 \leqq P(A) \leqq 1 \tag{5}$$

i.e. a probability is between 0 and 1.

***Theorem 1-16*:**

$$P(\emptyset) = 0 \tag{6}$$

i.e. the impossible event has probability zero.

***Theorem 1-17*:** If A' is the complement of A then

$$P(A') = 1 - P(A) \tag{7}$$

***Theorem 1-18*:** If $A = A_1 \cup A_2 \cup \cdots \cup A_n$, where $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{8}$$

In particular if $A = \mathscr{S}$, the sample space, then

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{9}$$

Theorem 1-19: If A and B are any two events then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{10}$$

More generally, if A_1, A_2, A_3 are any three events, then

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) \\ & - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) \\ & + P(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{11}$$

Generalizations to n events can also be made. See Problem 1.79.

Theorem 1-20: For any events A and B

$$P(A) = P(A \cap B) + P(A \cap B') \tag{12}$$

Theorem 1-21: If an event A must result in one of the mutually exclusive events $A_1, A_2, \ldots, A_n$ then

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + \cdots + P(A \cap A_n) \tag{13}$$

ASSIGNMENT OF PROBABILITIES

If a sample space $\mathscr{S}$ consists only of the elementary events $A_1, A_2, \ldots, A_n$ then by Theorem 1-18

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{14}$$

It follows that we can arbitrarily choose any nonnegative numbers for the probabilities of these simple events so long as (*14*) is satisfied. In particular, if we assume *equal probabilities* for all simple events, then

$$P(A_k) = \frac{1}{n} \qquad k = 1, 2, \ldots, n \tag{15}$$

and if A is any event made up of h such simple events we have

$$P(A) = \frac{h}{n} \tag{16}$$

This is equivalent to the classical approach to probability given on page 5. We could of course use other procedures for assigning probabilities, such as the frequency approach of page 6.

Assigning probabilities provides a *mathematical model* the success of which must be tested by experiment in much the same manner that theories in physics or other sciences must be tested by experiment.

EXAMPLE 1.21. A single die is tossed. Find the probability of a 2 or 5 turning up.

The sample space is $\mathscr{S} = \{1, 2, 3, 4, 5, 6\}$. If we assign equal probabilities to the sample points, i.e. if we assume that the die is fair, then

$$P(1) = P(2) = \cdots = P(6) = \frac{1}{6}$$

The event that either 2 or 5 turns up is indicated by $2 \cup 5$. Thus

$$P(2 \cup 5) = P(2) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

CONDITIONAL PROBABILITY

Let A and B be two events (Fig. 1-9) such that $P(A) > 0$. Denote by $P(B \mid A)$ the probability of B *given that* A has occurred. Since A is known to have occurred, it becomes the new sample space replacing the original $\mathcal{S}$. From this we are led to the definition

$$P(B \mid A) \equiv \frac{P(A \cap B)}{P(A)} \tag{17}$$

or

$$P(A \cap B) \equiv P(A)\,P(B \mid A) \tag{18}$$

In words, (*18*) says that the probability that both A and B occur is equal to the probability that A occurs times the probability that B occurs given that A has occurred. We call $P(B \mid A)$ the *conditional probability* of B given A, i.e. the probability that B will occur given that A has occurred. It is easy to show that conditional probability satisfies the axioms on page 6.

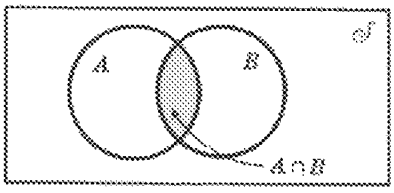


Fig. 1-9

EXAMPLE 1.22. Find the probability that a single toss of a die results in a number less than 4 if (*a*) no other information is given, (*b*) it is given that the toss resulted in an odd number.

(*a*) Let B denote the event {less than 4}. Since B is the union of the events 1, 2 or 3 turning up we see by Theorem 1-18 that

$$P(B) = P(1) + P(2) + P(3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

assuming equal probabilities for the sample points.

(*b*) Letting A be the event {odd number} we see that $P(A) = \frac{3}{6} = \frac{1}{2}$. Also $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$. Then

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Thus the added knowledge that the toss results in an odd number raises the probability from 1/2 to 2/3.

THEOREMS ON CONDITIONAL PROBABILITY

Theorem 1-22: For any three events A_1, A_2, A_3 we have

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \tag{19}$$

In words, the probability that A_1 and A_2 and A_3 all occur is equal to the probability that A_1 occurs times the probability that A_2 occurs given that A_1 has occurred times the probability that A_3 occurs given that both A_1 and A_2 have occurred. The result is easily generalized to n events.

Theorem 1-23: If an event A must result in one of the mutually exclusive events $A_1, A_2, \ldots, A_n$ then

$$P(A) = P(A_1)\,P(A \mid A_1) + P(A_2)\,P(A \mid A_2) + \cdots + P(A_n)\,P(A \mid A_n) \tag{20}$$

INDEPENDENT EVENTS

If $P(B \mid A) = P(B)$, i.e. the probability of B occurring is not affected by the occurrence or nonoccurrence of A, then we say that A and B are *independent events*. This is equivalent to

$$P(A \cap B) = P(A)\,P(B) \tag{21}$$

as seen from (*18*). Conversely, if (*21*) holds then A and B are independent. Some properties of independence are given in Problems 1.91 and 1.92.

If A_1, A_2, A_3 are to be independent then they must be pairwise independent,

$$P(A_j \cap A_k) = P(A_j)\,P(A_k) \qquad j \neq k \text{ where } j, k = 1, 2, 3 \tag{22}$$

and we must also have

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2)\,P(A_3) \tag{23}$$

Neither (*22*) nor (*23*) is by itself sufficient. Generalization to more than three events is easily made.

BAYES' THEOREM OR RULE

Suppose that $A_1, A_2, \ldots, A_n$ are mutually exclusive events whose union is the sample space $\mathcal{S}$, i.e. one of the events must occur. Then if A is any event we have the following important theorem:

Theorem 1-24 **(Bayes' rule):**

$$P(A_k \mid A) = \frac{P(A_k)\,P(A \mid A_k)}{\sum_{k=1}^{n} P(A_k)\,P(A \mid A_k)} \tag{24}$$

This enables us to find the probabilities of the various events $A_1, A_2, \ldots, A_n$ which can *cause* A to occur. For this reason Bayes' theorem is often referred to as a theorem on the *probability of causes.*

COMBINATORIAL ANALYSIS

In many cases the number of sample points in a sample space is not very large and so direct enumeration or counting of sample points needed to obtain probabilities is not difficult. However, problems arise where direct counting becomes a practical impossibility. In such cases use is made of *combinatorial analysis*, which could be called a *sophisticated way* of counting.

FUNDAMENTAL PRINCIPLE OF COUNTING. TREE DIAGRAMS

If one thing can be accomplished in n_1 different ways and after this a second thing can be accomplished in n_2 different ways, ..., and finally a kth thing can be accomplished in n_k different ways, then all k things can be accomplished in the specified order in $n_1 n_2 \cdots n_k$ different ways.

EXAMPLE 1.23. If a man has 2 shirts and 4 ties then he has $2 \cdot 4 = 8$ ways of choosing a shirt and then a tie.

A diagram, called a *tree diagram* because of its appearance (Fig. 1-10), is often used in connection with the above principle.

EXAMPLE 1.24. Letting the shirts be represented by S_1, S_2 and the ties by T_1, T_2, T_3, T_4, the various ways of choosing a shirt and then a tie are indicated in the tree diagram of Fig. 1-10.

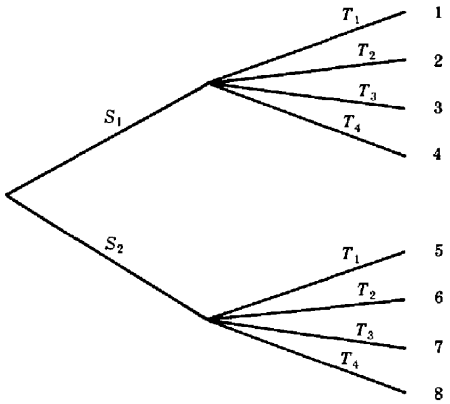


Fig. 1-10

PERMUTATIONS

Suppose that we are given n distinct objects and wish to *arrange* r of these objects in a line. Since there are n ways of choosing the 1st object, and after this is done $n-1$ ways of choosing the 2nd object, ..., and finally $n-r+1$ ways of choosing the rth object, it follows by the fundamental principle of counting that the number of different *arrangements*, or *permutations* as they are often called, is given by

$$ {}_nP_r = n(n-1)(n-2)\cdots(n-r+1) \tag{25} $$

where it is noted that the product has r factors. We call ${}_nP_r$ the number of *permutations of n objects taken r at a time.*

In the particular case where $r = n$, (25) becomes

$$ {}_nP_n = n(n-1)(n-2)\cdots 1 = n! \tag{26} $$

which is called n *factorial.* We can write (25) in terms of factorials as

$$ {}_nP_r = \frac{n!}{(n-r)!} \tag{27} $$

If $r = n$ we see that (27) and (26) agree only if we have $0! = 1$ and we shall actually take this as the definition of $0!$.

EXAMPLE 1.25. The number of different arrangements or permutations consisting of 3 letters each which can be formed from the 7 letters A, B, C, D, E, F, G is

$$ {}_7P_3 = \frac{7!}{4!} = 7 \cdot 6 \cdot 5 = 210 $$

Suppose that a set consists of n objects of which n_1 are of one type (i.e. indistinguishable from each other), n_2 are of a second type, ..., n_k are of a kth type. Here, of course, $n = n_1 + n_2 + \cdots + n_k$. Then the number of different permutations of the objects is

$$ {}_nP_{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!\, n_2! \cdots n_k!} \tag{28} $$

See Problem 1.34.

EXAMPLE 1.26. The number of different permutations of the 11 letters of the word $MISSISSIPPI$, which consists of 1 M, 4 I's, 4 S's and 2 P's, is

$$ \frac{11!}{1!\,4!\,4!\,2!} = 34{,}650 $$

COMBINATIONS

In a permutation we are interested in the order of arrangement of the objects. Thus abc is a different permutation from bca. In many problems, however, we are interested only in selecting or choosing objects without regard to order. Such selections are called *combinations.* For example abc and bca are the same combination.

Ziyad Asaad

The total number of combinations of r objects selected from n (also called the combinations of n *things taken* r *at a time*) is denoted by ${}_nC_r$ or $\binom{n}{r}$. We have (see Problem 1.36)

$$\binom{n}{r} = {}_nC_r = \frac{n!}{r!\,(n-r)!} \tag{29}$$

It can also be written

$$\binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{{}_nP_r}{r!} \tag{30}$$

It is easy to show that

$$\binom{n}{r} = \binom{n}{n-r} \quad \text{or} \quad {}_nC_r = {}_nC_{n-r} \tag{31}$$

EXAMPLE 1.27. The number of ways in which 3 cards can be chosen or selected from a total of 8 different cards is

$${}_8C_3 = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3!} = 56$$

BINOMIAL COEFFICIENTS

The numbers (29) are often called *binomial coefficients* because they arise in the *binomial expansion*

$$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n}y^n \tag{32}$$

They have many interesting properties.

EXAMPLE 1.28.

$$\begin{aligned}(x+y)^4 &= x^4 + \binom{4}{1}x^3y + \binom{4}{2}x^2y^2 + \binom{4}{3}xy^3 + \binom{4}{4}y^4 \\ &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4\end{aligned}$$

STIRLING'S APPROXIMATION TO $n!$

When n is large a direct evaluation of $n!$ is impractical. In such case use is made of the approximate formula

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n} \tag{33}$$

where $e = 2.71828\ldots$ is the base of natural logarithms. See Problem 1.48. The symbol $\sim$ in (33) means that the ratio of the left side to the right side approaches 1 as $n \to \infty$. For this reason we often call the right side an *asymptotic expansion* of the left side. For further discussion of Stirling's formula see Appendix A.

Solved Problems

SETS

1.1. Let A be the set of all real numbers whose squares are equal to 25. Show how to describe A by (*a*) the property method and (*b*) the roster method.

(*a*) $A = \{x \mid x^2 = 25\}$ which is read "the set of all elements x such that $x^2 = 25$."

(*b*) Since $x^2 = 25$ for $x = 5$ and $x = -5$, we can write $A = \{5, -5\}$, i.e. A is described by actually giving its elements.

1.2. Let $A = \{x \mid x \text{ is an odd integer}\}$, $B = \{x \mid x^2 - 8x + 15 = 0\}$. Show that $B \subset A$.

Since $x^2 - 8x + 15 = 0$ or $(x-3)(x-5) = 0$ if and only if $x = 3$ or $x = 5$, we have $B = \{3, 5\}$. Then since the elements 3 and 5 are both odd integers, they belong to A. Thus every element of B belongs to A and so $B \subset A$, i.e. B is a subset of A.

1.3. Is it true that $\{2\} = 2$?

No, 2 is a *real number* while $\{2\}$ is a *set* which consists of the real number 2. A set such as $\{2\}$ consisting of only one element is sometimes called a *singleton set* and must be distinguished from the element which it contains.

1.4. Determine which of the following statements are true and correct any which are false.

(*a*) $\{x \mid x \neq x\} = \{\emptyset\}$.

(*b*) If $A = \{x \mid x^2 = 4,\ x > 9\}$ and $B = \{x \mid x \leqq 1\}$, then $B \supset A$.

(*a*) The statement is false. Any particular object is assumed to be the same as, i.e. equal to, itself. Thus there is no object which is not equal to itself. Then $\{x \mid x \neq x\} = \emptyset$, the empty set.

The error lies in writing $\{\emptyset\}$ rather than $\emptyset$, since $\{\emptyset\}$ is a *nonempty* set which consists of the empty set.

(*b*) Note that this is read "A is the set of all x such that $x^2 = 4$ and $x > 9$."

Since there is no number x such that $x^2 = 4$ [or $x = 2, -2$] and $x > 9$, it follows that $A = \emptyset$. Also since the empty set is a subset of every set, it follows that $A \subset B$ or $B \supset A$ and the statement is true.

1.5. Prove that if $A \subset B$ and $B \subset C$, then $A \subset C$.

Let x be any element of A, i.e. $x \in A$. Then since $A \subset B$, i.e. every element of A is in B, we have $x \in B$. Also since $B \subset C$, we have $x \in C$. Thus every element of A is an element of C and so $A \subset C$.

SET OPERATIONS, VENN DIAGRAMS AND THEOREMS ON SETS

1.6. If the universe $\mathcal{U} = \{\frac{1}{2}, 0, \pi, 5, -\sqrt{2}, -4\}$ and subsets of $\mathcal{U}$ are given by $A = \{-\sqrt{2}, \pi, 0\}$, $B = \{5, \frac{1}{2}, -\sqrt{2}, -4\}$ and $C = \{\frac{1}{2}, -4\}$, find (*a*) $A \cap B$, (*b*) $A \cup B$, (*c*) $(A \cup B) \cap C$, (*d*) $B' \cup C'$, (*e*) $A - B$, (*f*) $(B \cap C)'$, (*g*) $(A \cap C) \cup (B \cap C)$.

(*a*) $A \cap B = \{-\sqrt{2}, \pi, 0\} \cap \{5, \frac{1}{2}, -\sqrt{2}, -4\} = \{-\sqrt{2}\}$

(*b*) $A \cup B = \{-\sqrt{2}, \pi, 0\} \cup \{5, \frac{1}{2}, -\sqrt{2}, -4\} = \{-\sqrt{2}, \pi, 0, 5, \frac{1}{2}, -4\}$

(*c*) $(A \cup B) \cap C = \{-\sqrt{2}, \pi, 0, 5, \frac{1}{2}, -4\} \cap \{\frac{1}{2}, -4\} = \{\frac{1}{2}, -4\}$ using the result of (*b*).

(*d*) B' = set of all elements in $\mathcal{U}$ which are not in $B = \{0, \pi\}$.

C' = set of all elements in $\mathcal{U}$ which are not in $C = \{0, \pi, 5, -\sqrt{2}\}$.

Then $B' \cup C' = \{0, \pi\} \cup \{0, \pi, 5, -\sqrt{2}\} = \{0, \pi, 5, -\sqrt{2}\}$.

(*e*) $A - B$ = set of elements in A which are not in $B = \{0, \pi\}$.

Another method. By Theorem 1-8, page 3, we have

$$A - B = A \cap B' = \{\tfrac{1}{2}, 0, \pi, 5, -\sqrt{2}, -4\} \cap \{0, \pi\} = \{0, \pi\}$$

(*f*) $B \cap C = \{5, \frac{1}{2}, -\sqrt{2}, -4\} \cap \{\frac{1}{2}, -4\} = \{\frac{1}{2}, -4\}$.

Then $(B \cap C)' = \{0, \pi, 5, -\sqrt{2}\}$.

Note that this result together with that of part (*d*) illustrates De Morgan's second law, Theorem 1-12*b*, page 3.

(*g*) $A \cap C = \{-\sqrt{2}, \pi, 0\} \cap \{\frac{1}{2}, -4\} = \emptyset$, the empty set.

$B \cap C = \{\frac{1}{2}, -4\}$ [see part (*f*)].

Then $$(A \cap C) \cup (B \cap C) = \emptyset \cup \{\tfrac{1}{2}, -4\} = \{\tfrac{1}{2}, -4\}$$

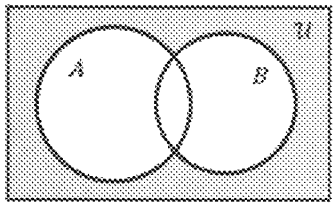
1.7. (*a*) Prove De Morgan's first law, Theorem 1-12*a*, page 3: $(A \cup B)' = A' \cap B'$. (*b*) Illustrate the result of part (*a*) by using a Venn diagram.

(*a*) We have

$$(A \cup B)' = \{x \mid x \notin A \cup B\} = \{x \mid x \notin A, x \notin B\} = \{x \mid x \in A', x \in B'\} = A' \cap B'$$

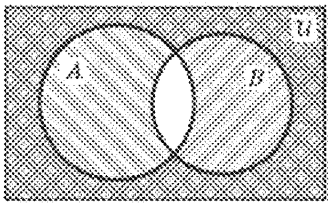
The result can be extended to any finite number of sets (see Problem 1.69).

(*b*) In the Venn diagram of Fig. 1-11 the shaded part represents $(A \cup B)'$. In Fig. 1-12, A' is indicated by parallel lines constructed from upper right to lower left while B' is indicated by parallel lines constructed from upper left to lower right. Then $A' \cap B'$ is represented by the cross hatched region where both sets of lines are present, and it is seen that this is the same as the shaded region of Fig. 1-11.



Shaded part = $(A \cup B)'$

Fig. 1-11



Cross hatched part = $A' \cap B'$

Fig. 1-12

Note that a Venn diagram does not provide a proof such as is given in part (*a*). However, it does serve to provide possible relationships among sets which can then be proved by methods similar to that given in (*a*).

1.8. Prove the first distributive law, Theorem 1-6, page 3: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

We have
$$\begin{aligned} A \cap (B \cup C) &= \{x \mid x \in A, x \in B \cup C\} \\ &= \{x \mid x \in A, x \in B \text{ or } x \in C\} \\ &= \{x \mid x \in A, x \in B \text{ or } x \in A, x \in C\} \\ &= \{x \mid x \in A \cap B \text{ or } x \in A \cap C\} \\ &= (A \cap B) \cup (A \cap C) \end{aligned}$$

1.9. Prove that for any sets A and B we have $A = (A \cap B) \cup (A \cap B')$.

Method 1.

$$A = \{x \mid x \in A\} = \{x \mid x \in A \cap B \text{ or } x \in A \cap B'\} = (A \cap B) \cup (A \cap B')$$

Method 2.

Let $C = B'$ in Problem 1.8. Then

$$\begin{aligned} A \cap (B \cup B') &= (A \cap B) \cup (A \cap B') \\ A \cap \mathcal{U} &= (A \cap B) \cup (A \cap B') \\ A &= (A \cap B) \cup (A \cap B') \end{aligned}$$

The result can be generalized (see Problem 1.74).

RANDOM EXPERIMENTS, SAMPLE SPACES AND EVENTS

1.10. A card is drawn at random from an ordinary deck of 52 playing cards. Describe the sample space if consideration of suits (*a*) is not, (*b*) is, taken into account.

(*a*) If we do not take into account the suits the sample space consists of ace, two, ..., ten, jack, queen, king, and it can be indicated as {1, 2, ..., 13}.

(*b*) If we do take into account the suits the sample space consists of ace of hearts, spades, diamonds and clubs; ...; king of hearts, spades, diamonds and clubs. Denoting hearts, spades, diamonds and clubs respectively by 1, 2, 3, 4, for example, we can indicate a jack of spades by (11, 2). The sample space then consists of the 52 points shown in Fig. 1-13.

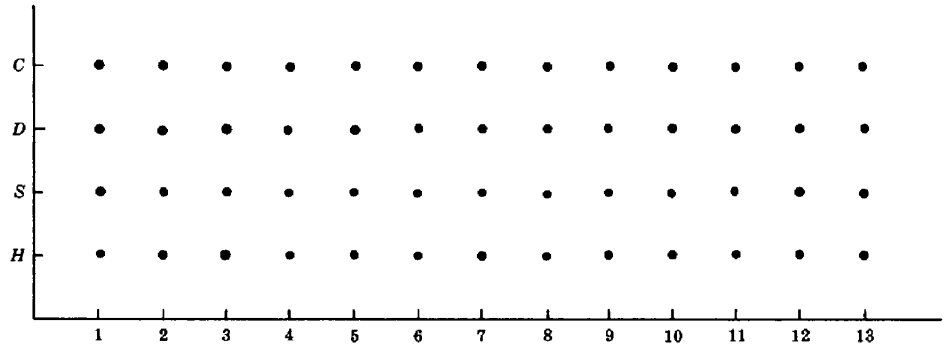


Fig. 1-13

1.11. Referring to the experiment of Problem 1.10 let A be the event {king is drawn} or simply {king} and B the event {club is drawn} or simply {club}. Describe the events (*a*) $A \cup B$, (*b*) $A \cap B$, (*c*) $A \cup B'$, (*d*) $A' \cup B'$, (*e*) $A - B$, (*f*) $A' - B'$, (*g*) $(A \cap B) \cup (A \cap B')$.

(*a*) $A \cup B$ = {either king or club (or both, i.e. king of clubs)}

(*b*) $A \cap B$ = {both king and club} = {king of clubs}

(*c*) Since B = {club}, B' = {not club} = {heart, diamond, spade}.

Then $A \cup B'$ = {king or heart or diamond or spade}.

(*d*) $A' \cup B'$ = {not king or not club} = {not king of clubs} = {any card but king of clubs}

This can also be seen by noting that $A' \cup B' = (A \cap B)'$ and using (*b*).

(*e*) $A - B$ = {king but not club}

This is the same as $A \cap B'$ = {king and not club}.

(*f*) $A' - B'$ = {not king and not "not club"} = {not king and club} = {any club except king}.

This can also be seen by noting that $A' - B' = A' \cap (B')' = A' \cap B$.

(*g*) $(A \cap B) \cup (A \cap B')$ = {(king and club) or (king and not club)} = {king}.

This can also be seen by noting that $(A \cap B) \cup (A \cap B') = A$.

1.12. Use Fig. 1-13 to describe the events (*a*) $A \cup B$, (*b*) $A' \cap B'$.

The required events are indicated by the Venn diagram of Fig. 1-14. In a similar manner all the events of Problem 1.11 can also be indicated by Venn diagrams. It should be observed from Fig. 1-14 that $A' \cap B'$ is the complement of $A \cup B$, in agreement with Theorem 1-12*a*, page 3.

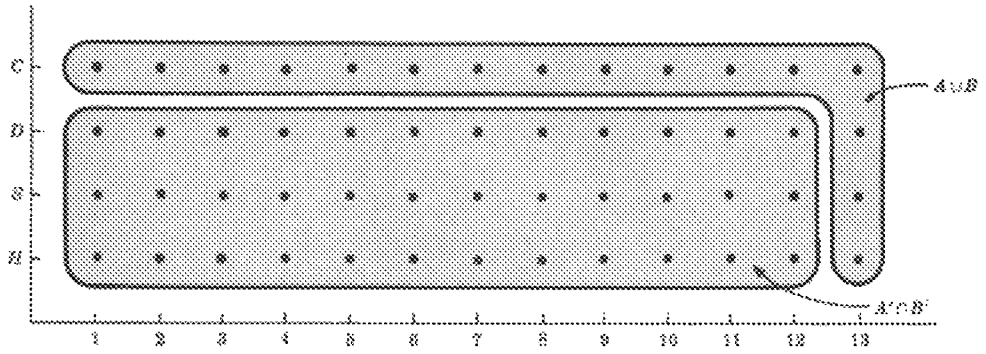


Fig. 1-14

THEOREMS ON PROBABILITY

1.13. Prove (*a*) Theorem 1-14, (*b*) Theorem 1-15, (*c*) Theorem 1-16, page 6.

(*a*) We have $A_2 = A_1 \cup (A_2 - A_1)$ where A_1 and $A_2 - A_1$ are mutually exclusive. Then by Axiom 3, page 6,

$$P(A_2) = P(A_1) + P(A_2 - A_1)$$

so that

$$P(A_2 - A_1) = P(A_2) - P(A_1)$$

Since $P(A_2 - A_1) \geq 0$ by Axiom 1, page 6, it also follows that $P(A_2) \geq P(A_1)$.

(*b*) We already know that $P(A) \geq 0$ by Axiom 1. To prove that $P(A) \leq 1$ we first note that $A \subset \mathcal{S}$. Thus by Theorem 1-14 [part (*a*)] and Axiom 2

$$P(A) \leq P(\mathcal{S}) = 1$$

(*c*) We have $\mathcal{S} = \mathcal{S} \cup \emptyset$. Since $\mathcal{S} \cap \emptyset = \emptyset$ it follows from Axiom 3 that

$$P(\mathcal{S}) = P(\mathcal{S}) + P(\emptyset) \quad \text{or} \quad P(\emptyset) = 0$$

1.14. Prove (*a*) Theorem 1-17 and (*b*) Theorem 1-19.

(*a*) We have $A \cup A' = \mathcal{S}$. Then since $A \cap A' = \emptyset$ we have

$$P(A \cup A') = P(\mathcal{S}) \quad \text{or} \quad P(A) + P(A') = 1$$

i.e.

$$P(A') = 1 - P(A)$$

(*b*) We have from the Venn diagram of Fig. 1-15

(*1*) $$A \cup B = A \cup [B - (A \cap B)]$$

Then since the sets A and $B - (A \cap B)$ are mutually exclusive, we have on using Axiom 3 and Theorem 1-14:

$$P(A \cup B) = P(A) + P(B - (A \cap B))$$
$$= P(A) + P(B) - P(A \cap B)$$

Although we have used the Venn diagram the result (*1*) can be established directly (see Problem 1.77).

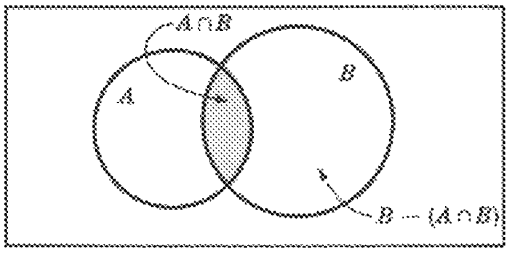


Fig. 1-15

CALCULATION OF PROBABILITIES

1.15. A card is drawn at random from an ordinary deck of 52 playing cards. Find the probability that it is (*a*) an ace, (*b*) a jack of hearts, (*c*) a three of clubs or a six of diamonds, (*d*) a heart, (*e*) any suit except hearts, (*f*) a ten or a spade, (*g*) neither a four nor a club.

Ziyad Asaad

Let us use for brevity H, S, D, C to indicate heart, spade, diamond, club, respectively, and $1, 2, \ldots, 13$ for ace, two, ..., king. Then $3 \cap H$ means three of hearts, while $3 \cup H$ means three or heart. Let us use the sample space of Problem 1.10(*b*), assigning equal probabilities of 1/52 to each sample point. Thus, for example, $P(6 \cap C) = 1/52$.

(*a*)
$$\begin{aligned} P(1) &= P(1 \cap H \text{ or } 1 \cap S \text{ or } 1 \cap D \text{ or } 1 \cap C) \\ &= P(1 \cap H) + P(1 \cap S) + P(1 \cap D) + P(1 \cap C) \\ &= \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{1}{13} \end{aligned}$$

This could also have been achieved from the sample space of Problem 1.10(*a*) where each sample point, in particular "ace," has probability 1/13. It could also have been arrived at by simply reasoning that there are 13 numbers and so each has probability 1/13 of being drawn.

(*b*) $P(11 \cap H) = \dfrac{1}{52}$

(*c*) $P(3 \cap C \text{ or } 6 \cap D) = P(3 \cap C) + P(6 \cap D) = \dfrac{1}{52} + \dfrac{1}{52} = \dfrac{1}{26}$

(*d*) $P(H) = P(1 \cap H \text{ or } 2 \cap H \text{ or } \ldots 13 \cap H) = \dfrac{1}{52} + \dfrac{1}{52} + \cdots + \dfrac{1}{52} = \dfrac{13}{52} = \dfrac{1}{4}$

This could also have been arrived at by noting that there are four suits and each has equal probability $\frac{1}{4}$ of being drawn.

(*e*) $P(H') = 1 - P(H) = 1 - \dfrac{1}{4} = \dfrac{3}{4}$ using part (*d*) and Theorem 1-17, page 6.

(*f*) Since 10 and S are not mutually exclusive we have from Theorem 1-19

$$P(10 \cup S) = P(10) + P(S) - P(10 \cap S) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

(*g*) The probability of neither four nor club can be denoted by $P(4' \cap C')$. But by Theorem 1-12(*a*), page 3, $4' \cap C' = (4 \cup C)'$. Thus

$$\begin{aligned} P(4' \cap C') = P[(4 \cup C)'] &= 1 - P(4 \cup C) \\ &= 1 - [P(4) + P(C) - P(4 \cap C)] \\ &= 1 - \left[\frac{1}{13} + \frac{1}{4} - \frac{1}{52}\right] = \frac{9}{13} \end{aligned}$$

We could also get this by noting that the Venn diagram favorable to this event is the complement of the event shown shaded in Fig. 1-16. Since this complement has $52 - 16 = 36$ sample points in it and each sample point is assigned probability 1/52, the required probability is $36/52 = 9/13$.

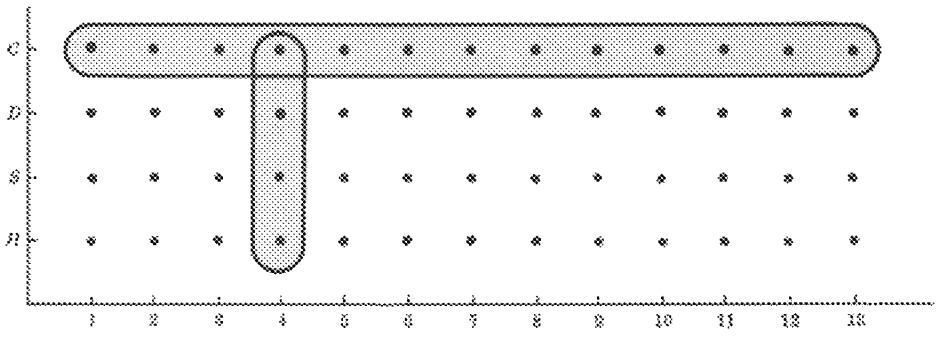


Fig. 1-16

1.16. A ball is drawn at random from a box containing 6 red balls, 4 white balls and 5 blue balls. Determine the probability that it is (*a*) red, (*b*) white, (*c*) blue, (*d*) not red, (*e*) red or white.

(*a*) **Method 1.**

Let R, W and B denote the events of drawing a red ball, white ball and blue ball, respectively. Then

$$P(R) = \frac{\text{ways of choosing a red ball}}{\text{total ways of choosing a ball}} = \frac{6}{6+4+5} = \frac{6}{15} = \frac{2}{5}$$

Method 2.

Our sample space consists of $6+4+5=15$ sample points. Then if we assign equal probabilities 1/15 to each sample point we see that $P(R) = 6/15 = 2/5$, since there are 6 sample points corresponding to "red ball."

(*b*) $P(W) = \dfrac{4}{6+4+5} = \dfrac{4}{15}$

(*c*) $P(B) = \dfrac{5}{6+4+5} = \dfrac{5}{15} = \dfrac{1}{3}$

(*d*) $P(\text{not red}) = P(R') = 1 - P(R) = 1 - \dfrac{2}{5} = \dfrac{3}{5}$ by part (*a*).

(*e*) **Method 1.**

$$P(\text{red or white}) = P(R \cup W) = \frac{\text{ways of choosing a red or white ball}}{\text{total ways of choosing a ball}}$$

$$= \frac{6+4}{6+4+5} = \frac{10}{15} = \frac{2}{3}$$

This can also be worked using the sample space as in part (*a*).

Method 2.

$$P(R \cup W) = P(B') = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3} \text{ by part } (e).$$

Method 3.

Since events R and W are mutually exclusive it follows from (*4*), page 6, that

$$P(R \cup W) = P(R) + P(W) = \frac{2}{5} + \frac{4}{15} = \frac{2}{3}$$

CONDITIONAL PROBABILITY AND INDEPENDENT EVENTS

1.17. A fair die is tossed twice. Find the probability of getting a 4, 5 or 6 on the first toss and a 1, 2, 3 or 4 on the second toss.

Let A_1 be the event "4, 5 or 6 on first toss," and A_2 be the event "1, 2, 3 or 4 on second toss." Then we are looking for $P(A_1 \cap A_2)$.

Method 1.

$$P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1) = P(A_1)\,P(A_2) = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = \frac{1}{3}$$

We have used here the fact that the result of the second toss is *independent* of the first so that $P(A_2 \mid A_1) = P(A_2)$. Also we have used $P(A_1) = 3/6$ (since 4, 5 or 6 are 3 out of 6 equally likely possibilities) and $P(A_2) = 4/6$ (since 1, 2, 3 or 4 are 4 out of 6 equally likely possibilities).

Method 2.

Each of the 6 ways in which a die can fall on the first toss can be associated with each of the 6 ways in which it can fall on the second toss, a total of $6 \cdot 6 = 36$ ways, all equally likely.

Each of the 3 ways in which A_1 can occur can be associated with each of the 4 ways in which A_2 can occur to give $3 \cdot 4 = 12$ ways in which both A_1 and A_2 can occur. Then

$$P(A_1 \cap A_2) = \frac{12}{36} = \frac{1}{3}$$

This shows directly that A_1 and A_2 are independent since

$$P(A_1 \cap A_2) = \frac{1}{3} = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = P(A_1)\,P(A_2)$$

1.18. Find the probability of not getting a 7 or 11 total on either of two tosses of a pair of fair dice.

The sample space for each toss of the dice is shown in Fig. 1-17. For example (5, 2) means that 5 comes up on the first die and 2 on the second. Since the dice are fair and there are 36 sample points we assign probability 1/36 to each.

If we let A be the event "7 or 11" then A is indicated by the shaded portion in Fig. 1-17. Since 8 points are included we have $P(A) = 8/36 = 2/9$. It follows that the probability of no 7 or 11 is given by

$$P(A') = 1 - P(A) = 1 - \frac{2}{9} = \frac{7}{9}$$

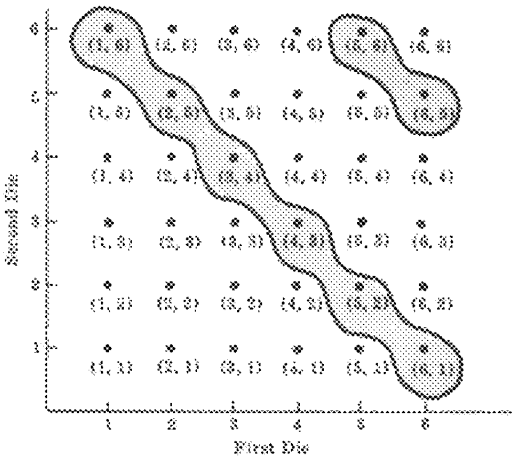


Fig. 1-17

Using subscripts 1, 2 to denote 1st and 2nd tosses of the dice we see that the probability of no 7 or 11 on either the first or second tosses is given by

$$P(A_1')\,P(A_2' \mid A_1') = P(A_1')\,P(A_2') = \left(\frac{7}{9}\right)\left(\frac{7}{9}\right) = \frac{49}{81}$$

using the fact that the tosses are independent.

1.19. Two cards are drawn from a well-shuffled ordinary deck of 52 cards. Find the probability that they are both aces if the first card is (*a*) replaced, (*b*) not replaced.

Method 1.

Let A_1 = event "ace on first draw" and A_2 = event "ace on second draw." Then we are looking for $P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1)$.

(*a*) Since for the first drawing there are 4 aces in 52 cards, $P(A_1) = 4/52$. Also, if the card is replaced for the second drawing, then $P(A_2 \mid A_1) = 4/52$, since there are also 4 aces out of 52 cards for the second drawing. Then

$$P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{4}{52}\right) = \frac{1}{169}$$

(*b*) As in part (*a*), $P(A_1) = 4/52$. However if an ace occurs on the first drawing there will be only 3 aces left in the remaining 51 cards, so that $P(A_2 \mid A_1) = 3/51$. Then

$$P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{1}{221}$$

Method 2.

(*a*) The first card can be drawn in any one of 52 ways and since there is replacement the second card can also be drawn in any one of 52 ways. Then both cards can be drawn in (52)(52) ways, all equally likely.

In such case there are 4 ways of choosing an ace on the first draw and 4 ways of choosing an ace on the second draw so that the number of ways of choosing aces on the first and second draws is (4)(4). Then the required probability is

$$\frac{(4)(4)}{(52)(52)} = \frac{1}{169}$$

(*b*) The first card can be drawn in any one of 52 ways and since there is no replacement the second card can be drawn in any one of 51 ways. Then both cards can be drawn in (52)(51) ways, all equally likely.

Ziyad Asaad

In such case there are 4 ways of choosing an ace on the first draw and 3 ways of choosing an ace on the second draw so that the number of ways of choosing aces on the first and second draws is (4)(3). Then the required probability is

$$\frac{(4)(3)}{(52)(51)} = \frac{1}{221}$$

1.20. Three balls are drawn successively from the box of Problem 1.16. Find the probability that they are drawn in the order red, white and blue if each ball is (*a*) replaced, (*b*) not replaced.

Let R_1 = event "red on first draw," W_2 = event "white on second draw," B_3 = event "blue on third draw." We require $P(R_1 \cap W_2 \cap B_3)$.

(*a*) If each ball is replaced, then the events are independent and

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(R_1)\,P(W_2 \mid R_1)\,P(B_3 \mid R_1 \cap W_2) \\ &= P(R_1)\,P(W_2)\,P(B_3) \\ &= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{6+4+5}\right)\left(\frac{5}{6+4+5}\right) = \frac{8}{225} \end{aligned}$$

(*b*) If each ball is not replaced, then the events are dependent and

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(R_1)\,P(W_2 \mid R_1)\,P(B_3 \mid R_1 \cap W_2) \\ &= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{5+4+5}\right)\left(\frac{5}{5+3+5}\right) = \frac{4}{91} \end{aligned}$$

1.21. Find the probability of a 4 turning up at least once in two tosses of a fair die.

Let A_1 = event "4 on first toss" and A_2 = event "4 on second toss." Then

$A_1 \cup A_2$ = event "4 on first toss or 4 on second toss or both"

= event "at least one 4 turns up"

and we require $P(A_1 \cup A_2)$.

Method 1.

Events A_1 and A_2 are not mutually exclusive, but they are independent. Hence, by (*10*) and (*21*),

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \\ &= P(A_1) + P(A_2) - P(A_1)\,P(A_2) \\ &= \frac{1}{6} + \frac{1}{6} - \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{11}{36} \end{aligned}$$

Method 2.

$$P(\text{at least one 4 comes up}) + P(\text{no 4 comes up}) = 1$$

Then

$$\begin{aligned} P(\text{at least one 4 comes up}) &= 1 - P(\text{no 4 comes up}) \\ &= 1 - P(\text{no 4 on 1st toss and no 4 on 2nd toss}) \\ &= 1 - P(A_1' \cap A_2') = 1 - P(A_1')\,P(A_2') \\ &= 1 - \left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{11}{36} \end{aligned}$$

Method 3.

Total number of equally likely ways in which both dice can fall $= 6 \cdot 6 = 36$.

Also, number of ways in which A_1 occurs but not A_2 = 5

number of ways in which A_2 occurs but not A_1 = 5

number of ways in which both A_1 and A_2 occur = 1

Then the number of ways in which at least one of the events A_1 or A_2 occurs $= 5 + 5 + 1 = 11$. Thus $P(A_1 \cup A_2) = 11/36$.

1.22. One bag contains 4 white balls and 2 black balls; another contains 3 white balls and 5 black balls. If one ball is drawn from each bag, find the probability that (*a*) both are white, (*b*) both are black, (*c*) one is white and one is black.

Let W_1 = event "white ball from first bag," W_2 = event "white ball from second bag."

(*a*) $$P(W_1 \cap W_2) = P(W_1)\,P(W_2 \mid W_1) = P(W_1)\,P(W_2) = \left(\frac{4}{4+2}\right)\left(\frac{3}{3+5}\right) = \frac{1}{4}$$

(*b*) $$P(W_1' \cap W_2') = P(W_1')\,P(W_2' \mid W_1') = P(W_1')\,P(W_2') = \left(\frac{2}{4+2}\right)\left(\frac{5}{3+5}\right) = \frac{5}{24}$$

(*c*) The required probability is

$$1 - P(W_1 \cap W_2) - P(W_1' \cap W_2') = 1 - \frac{1}{4} - \frac{5}{24} = \frac{13}{24}$$

1.23. Prove Theorem 1-23, page 8.

We prove the theorem for the case $n = 2$. Extensions to larger values of n are easily made.

If event A must result in one of the two mutually exclusive events A_1, A_2, then

$$A = (A \cap A_1) \cup (A \cap A_2)$$

But $A \cap A_1$ and $A \cap A_2$ are mutually exclusive since A_1 and A_2 are. Thus by Axiom 3

$$\begin{aligned} P(A) &= P(A \cap A_1) + P(A \cap A_2) \\ &= P(A_1)\,P(A \mid A_1) + P(A_2)\,P(A \mid A_2) \end{aligned}$$

using (*18*), page 8.

1.24. Box *I* contains 3 red and 2 blue marbles while Box *II* contains 2 red and 8 blue marbles. A fair coin is tossed. If the coin turns up heads a marble is chosen from Box *I*; if it turns up tails a marble is chosen from Box *II*. Find the probability that a red marble is chosen.

Let R denote the event "a red marble is chosen" while I and II denote the events that Box I and Box II are chosen, respectively. Since a red marble can result by choosing either Box I or II we can use the results of Problem 1.23 with $A = R$, $A_1 = I$, $A_2 = II$. Thus the probability of choosing a red marble is

$$P(R) = P(I)\,P(R \mid I) + P(II)\,P(R \mid II) = \left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right) = \frac{2}{5}$$

BAYES' THEOREM

1.25. Prove Bayes' theorem (Theorem 1-24, page 9).

Since A results in one of the mutually exclusive events $A_1, A_2, \ldots, A_n$ we have by Theorem 1-23 (Problem 1.23)

$$P(A) = P(A_1)\,P(A \mid A_1) + \cdots + P(A_n)\,P(A \mid A_n) = \sum_{k=1}^{n} P(A_k)\,P(A \mid A_k)$$

Thus $$P(A_k \mid A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A_k)\,P(A \mid A_k)}{\sum_{k=1}^{n} P(A_k)\,P(A \mid A_k)}$$

1.26. Suppose in Problem 1.24 that the one who tosses the coin does not reveal whether it has turned up heads or tails (so that the box from which a marble was chosen is not

revealed), but does reveal that a red marble was chosen. What is the probability that Box I was chosen (i.e. the coin turned up heads)?

Let us use the same terminology as in Problem 1.24, i.e. $A = R$, $A_1 = I$, $A_2 = II$. We seek the probability that Box I was chosen given that a red marble is known to have been chosen. Using Bayes' rule with $n = 2$ this probability is given by

$$P(I \mid R) = \frac{P(I)\,P(R \mid I)}{P(I)\,P(R \mid I) + P(II)\,P(R \mid II)} = \frac{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right)}{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right)} = \frac{3}{4}$$

Thus we might be willing to give 3 to 1 odds that Box I was chosen.

COMBINATORIAL ANALYSIS, COUNTING AND TREE DIAGRAMS

1.27. A committee of 3 members is to be formed consisting of one representative each from labor, management and the public. If there are 3 possible representatives from labor, 2 from management and 4 from the public, determine how many different committees can be formed, using (*a*) the fundamental principle of counting and (*b*) a tree diagram.

(*a*) We can choose a labor representative in 3 different ways and after this a management representative in 2 different ways. Then there are $3 \cdot 2 = 6$ different ways of choosing a labor and management representative. With each of these ways we can choose a public representative in 4 different ways. Thus the number of different committees which can be formed is $3 \cdot 2 \cdot 4 = 24$.

(*b*) Denote the 3 labor representatives by L_1, L_2, L_3; the management representatives by M_1, M_2; and the public representatives by P_1, P_2, P_3, P_4. Then the tree diagram of Fig. 1-18 shows that there are 24 different committees in all. From this tree diagram we can list all these different committees, e.g. $L_1M_1P_1$, $L_1M_1P_2$, etc.

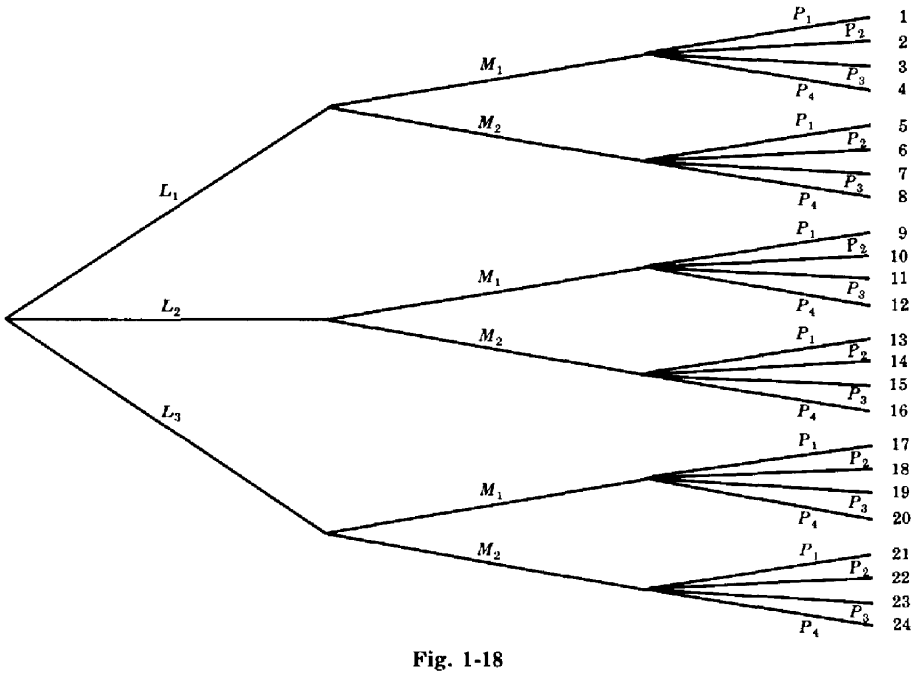


Fig. 1-18

PERMUTATIONS

1.28. In how many ways can 5 differently colored marbles be arranged in a row?

We must arrange the 5 marbles in 5 positions thus: — — — — —. The first position can be occupied by any one of five marbles, i.e. there are 5 ways of filling the first position. When this

has been done there are 4 ways of filling the second position. Then there are 3 ways of filling the third position, 2 ways of filling the fourth position, and finally only 1 way of filling the last position. Therefore:

$$\text{Number of arrangements of 5 marbles in a row} = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$$

In general,

$$\text{Number of arrangements of } n \text{ different objects in a row} = n(n-1)(n-2)\cdots 1 = n!$$

This is also called the number of *permutations* of n different objects taken n at a time and is denoted by ${}_nP_n$.

1.29. In how many ways can 10 people be seated on a bench if only 4 seats are available?

The first seat can be filled in any one of 10 ways and when this has been done there are 9 ways of filling the second seat, 8 ways of filling the third seat and 7 ways of filling the fourth seat. Therefore:

$$\text{Number of arrangements of 10 people taken 4 at a time} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

In general,

$$\text{Number of arrangements of } n \text{ different objects taken } r \text{ at a time} = n(n-1)\cdots(n-r+1)$$

This is also called the number of *permutations* of n different objects taken r at a time and is denoted by ${}_nP_r$. Note that when $r = n$, ${}_nP_n = n!$ as in Problem 1.28.

1.30. Evaluate (*a*) ${}_8P_3$, (*b*) ${}_6P_4$, (*c*) ${}_{15}P_1$, (*d*) ${}_3P_3$.

(*a*) ${}_8P_3 = 8 \cdot 7 \cdot 6 = 336$ (*b*) ${}_6P_4 = 6 \cdot 5 \cdot 4 \cdot 3 = 360$ (*c*) ${}_{15}P_1 = 15$ (*d*) ${}_3P_3 = 3 \cdot 2 \cdot 1 = 6$

1.31. It is required to seat 5 men and 4 women in a row so that the women occupy the even places. How many such arrangements are possible?

The men may be seated in ${}_5P_5$ ways and the women in ${}_4P_4$ ways. Each arrangement of the men may be associated with each arrangement of the women. Hence

$$\text{Number of arrangements} = {}_5P_5 \cdot {}_4P_4 = 5!\,4! = (120)(24) = 2880$$

1.32. How many 4-digit numbers can be formed with the 10 digits 0, 1, 2, 3, . . . , 9 if (*a*) repetitions are allowed, (*b*) repetitions are not allowed, (*c*) the last digit must be zero and repetitions are not allowed?

(*a*) The first digit can be any one of 9 (since 0 is not allowed). The second, third and fourth digits can be any one of 10. Then $9 \cdot 10 \cdot 10 \cdot 10 = 9000$ numbers can be formed.

(*b*) The first digit can be any one of 9 (any one but 0).

The second digit can be any one of 9 (any but that used for the first digit).

The third digit can be any one of 8 (any but those used for the first two digits).

The fourth digit can be any one of 7 (any but those used for the first three digits).

Then $9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

Another method.

The first digit can be any one of 9 and the remaining three can be chosen in ${}_9P_3$ ways. Then $9 \cdot {}_9P_3 = 9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

(*c*) The first digit can be chosen in 9 ways, the second in 8 ways and the third in 7 ways. Then $9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

Another method.

The first digit can be chosen in 9 ways and the next two digits in ${}_8P_2$ ways. Then $9 \cdot {}_8P_2 = 9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

1.33. Four different mathematics books, six different physics books, and two different chemistry books are to be arranged on a shelf. How many different arrangements are possible if (*a*) the books in each particular subject must all stand together, (*b*) only the mathematics books must stand together?

(*a*) The mathematics books can be arranged among themselves in ${}_4P_4 = 4!$ ways, the physics books in ${}_6P_6 = 6!$ ways, the chemistry books in ${}_2P_2 = 2!$ ways, and the three groups in ${}_3P_3 = 3!$ ways. Thus

$$\text{Number of arrangements} = 4!\,6!\,2!\,3! = 207{,}360$$

(*b*) Consider the four mathematics books as one big book. Then we have 9 books which can be arranged in ${}_9P_9 = 9!$ ways. In all of these ways the mathematics books are together. But the mathematics books can be arranged among themselves in ${}_4P_4 = 4!$ ways. Hence

$$\text{Number of arrangements} = 9!\,4! = 8{,}709{,}120$$

1.34. Five red marbles, two white marbles, and three blue marbles are arranged in a row. If all the marbles of the same color are not distinguishable from each other, how many different arrangements are possible?

Assume that there are N different arrangements. Multiplying N by the numbers of ways of arranging (*a*) the five red marbles among themselves, (*b*) the two white marbles among themselves, and (*c*) the three blue marbles among themselves (i.e. multiplying N by $5!\,2!\,3!$), we obtain the number of ways of arranging the 10 marbles if they were all distinguishable, i.e. $10!$.

Then $(5!\,2!\,3!)N = 10!$ and $N = 10!/(5!\,2!\,3!)$.

In general, the number of different arrangements of n objects of which n_1 are alike, n_2 are alike, ..., n_k are alike is $\dfrac{n!}{n_1!\,n_2!\cdots n_k!}$ where $n_1 + n_2 + \cdots + n_k = n$.

1.35. In how many ways can 7 people be seated at a round table if (*a*) they can sit anywhere, (*b*) 2 particular people must not sit next to each other?

(*a*) Let 1 of them be seated anywhere. Then the remaining 6 people can be seated in $6! = 720$ ways, which is the total number of ways of arranging the 7 people in a circle.

(*b*) Consider the 2 particular people as one person. Then there are 6 people altogether and they can be arranged in 5! ways. But the 2 people considered as 1 can be arranged among themselves in 2! ways. Thus the number of ways of arranging 7 people at a round table with 2 particular people sitting together $= 5!\,2! = 240$.

Then using (*a*), the total number of ways in which 7 people can be seated at a round table so that the 2 particular people do not sit together $= 720 - 240 = 480$ ways.

COMBINATIONS

1.36. In how many ways can 10 objects be split into two groups containing 4 and 6 objects respectively?

This is the same as the number of arrangements of 10 objects of which 4 objects are alike and 6 other objects are alike. By Problem 1.34 this is $\dfrac{10!}{4!\,6!} = \dfrac{10\cdot 9\cdot 8\cdot 7}{4!} = 210$.

The problem is equivalent to finding the number of selections of 4 out of 10 objects (or 6 out of 10 objects), the order of selection being immaterial. In general the number of selections of r out of n objects, called the number of *combinations* of n things taken r at a time, is denoted by ${}_nC_r$ or $\binom{n}{r}$ and is given by

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{{}_nP_r}{r!}$$

1.37. Evaluate (*a*) ${}_7C_4$, (*b*) ${}_6C_5$, (*c*) ${}_4C_4$.

(*a*) ${}_7C_4 = \dfrac{7!}{4!\,3!} = \dfrac{7\cdot6\cdot5\cdot4}{4!} = \dfrac{7\cdot6\cdot5}{3\cdot2\cdot1} = 35.$

(*b*) ${}_6C_5 = \dfrac{6!}{5!\,1!} = \dfrac{6\cdot5\cdot4\cdot3\cdot2}{5!} = 6,$ or ${}_6C_5 = {}_6C_1 = 6.$

(*c*) ${}_4C_4$ is the number of selections of 4 objects taken 4 at a time, and there is only one such selection. Then ${}_4C_4 = 1$.

Note that formally ${}_4C_4 = \dfrac{4!}{4!\,0!} = 1$ if we *define* $0! = 1$.

1.38. In how many ways can a committee of 5 people be chosen out of 9 people?

$$\binom{9}{5} = {}_9C_5 = \frac{9!}{5!\,4!} = \frac{9\cdot8\cdot7\cdot6\cdot5}{5!} = 126$$

1.39. Out of 5 mathematicians and 7 physicists, a committee consisting of 2 mathematicians and 3 physicists is to be formed. In how many ways can this be done if (*a*) any mathematician and any physicist can be included, (*b*) one particular physicist must be on the committee, (*c*) two particular mathematicians cannot be on the committee?

(*a*) 2 mathematicians out of 5 can be selected in ${}_5C_2$ ways.

3 physicists out of 7 can be selected in ${}_7C_3$ ways.

Total number of possible selections $= {}_5C_2 \cdot {}_7C_3 = 10\cdot35 = 350$

(*b*) 2 mathematicians out of 5 can be selected in ${}_5C_2$ ways.

2 physicists out of 6 can be selected in ${}_6C_2$ ways.

Total number of possible selections $= {}_5C_2 \cdot {}_6C_2 = 10\cdot15 = 150$

(*c*) 2 mathematicians out of 3 can be selected in ${}_3C_2$ ways.
3 physicists out of 7 can be selected in ${}_7C_3$ ways.

Total number of possible selections $= {}_3C_2 \cdot {}_7C_3 = 3\cdot35 = 105$

1.40. How many different salads can be made from lettuce, escarole, endive, watercress and chicory?

Each green can be dealt with in 2 ways, as it can be chosen or not chosen. Since each of the 2 ways of dealing with a green is associated with 2 ways of dealing with each of the other greens, the number of ways of dealing with the five greens $= 2^5$ ways. But 2^5 ways includes the case in which no green is chosen. Hence

Number of salads $= 2^5 - 1 = 31$

Another method.

One can select either 1 out of 5 greens, 2 out of 5 greens, ..., 5 out of 5 greens. Then the required number of salads is

$${}_5C_1 + {}_5C_2 + {}_5C_3 + {}_5C_4 + {}_5C_5 = 5 + 10 + 10 + 5 + 1 = 31$$

In general, for any positive integer n, ${}_nC_1 + {}_nC_2 + {}_nC_3 + \cdots + {}_nC_n = 2^n - 1$.

1.41. From 7 consonants and 5 vowels, how many words can be formed consisting of 4 different consonants and 3 different vowels? The words need not have meaning.

The 4 different consonants can be selected in ${}_7C_4$ ways, the 3 different vowels can be selected in ${}_5C_3$ ways, and the resulting 7 different letters (4 consonants, 3 vowels) can then be arranged among themselves in ${}_7P_7 = 7!$ ways. Then

Number of words $= {}_7C_4 \cdot {}_5C_3 \cdot 7! = 35\cdot10\cdot5040 = 1{,}764{,}000$

THE BINOMIAL COEFFICIENTS

1.42. Prove that $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

We have

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!} = \frac{n(n-1)!}{r!\,(n-r)!} = \frac{(n-r+r)(n-1)!}{r!\,(n-r)!}$$

$$= \frac{(n-r)(n-1)!}{r!\,(n-r)!} + \frac{r(n-1)!}{r!\,(n-r)!}$$

$$= \frac{(n-1)!}{r!\,(n-r-1)!} + \frac{(n-1)!}{(r-1)!\,(n-r)!}$$

$$= \binom{n-1}{r} + \binom{n-1}{r-1}$$

The result has the following interesting application. If we write out the coefficients in the binomial expansion of $(x+y)^n$ for $n = 0, 1, 2, \ldots$ we obtain the following arrangement, called *Pascal's triangle*:

$n = 0$ 1

$n = 1$ 1 1

$n = 2$ 1 2 1

$n = 3$ 1 3 3 1

$n = 4$ 1 4 6 4 1

$n = 5$ 1 5 10 10 5 1

$n = 6$ 1 6 15 20 15 6 1

etc.

An entry in any line can be obtained by adding the two entries in the preceding line which are to its immediate left and right. Thus $10 = 4 + 6$, $15 = 10 + 5$, etc.

1.43. Find the constant term in the expansion of $\left(x^2 + \frac{1}{x}\right)^{12}$.

According to the binomial theorem

$$\left(x^2 + \frac{1}{x}\right)^{12} = \sum_{k=0}^{12} \binom{12}{k} (x^2)^k \left(\frac{1}{x}\right)^{12-k} = \sum_{k=0}^{12} \binom{12}{k} x^{3k-12}$$

The constant term corresponds to the one for which $3k - 12 = 0$, i.e. $k = 4$, and is thus given by

$$\binom{12}{4} = \frac{12 \cdot 11 \cdot 10 \cdot 9}{4 \cdot 3 \cdot 2 \cdot 1} = 495$$

PROBABILITY USING COMBINATORIAL ANALYSIS

1.44. A box contains 8 red, 3 white and 9 blue balls. If 3 balls are drawn at random without replacement, determine the probability that (*a*) all 3 are red, (*b*) all 3 are white, (*c*) 2 are red and 1 is white, (*d*) at least 1 is white, (*e*) 1 of each color is drawn, (*f*) the balls are drawn in the order red, white, blue.

(*a*) **Method 1.**

Let R_1, R_2, R_3 denote the events, "red ball on 1st draw," "red ball on 2nd draw," "red ball on 3rd draw," respectively. Then $R_1 \cap R_2 \cap R_3$ denotes the event "all 3 balls drawn are red." We thus have

$$P(R_1 \cap R_2 \cap R_3) = P(R_1)\,P(R_2 \mid R_1)\,P(R_3 \mid R_1 \cap R_2)$$
$$= \left(\frac{8}{20}\right)\left(\frac{7}{19}\right)\left(\frac{6}{18}\right) = \frac{14}{285}$$

Method 2.

$$\text{Required probability} = \frac{\text{number of selections of 3 out of 8 red balls}}{\text{number of selections of 3 out of 20 balls}} = \frac{{}_8C_3}{{}_{20}C_3} = \frac{14}{285}$$

(*b*) Using the second method indicated in part (*a*),

$$P(\text{all 3 are white}) = \frac{{}_3C_3}{{}_{20}C_3} = \frac{1}{1140}$$

The first method indicated in part (*a*) can also be used.

(*c*)
$$P(\text{2 are red and 1 is white})$$
$$= \frac{(\text{selections of 2 out of 8 red balls})(\text{selections of 1 out of 3 white balls})}{\text{number of selections of 3 out of 20 balls}}$$
$$= \frac{({}_8C_2)({}_3C_1)}{{}_{20}C_3} = \frac{7}{95}$$

(*d*) $P(\text{none is white}) = \dfrac{{}_{17}C_3}{{}_{20}C_3} = \dfrac{34}{57}$. Then

$$P(\text{at least 1 is white}) = 1 - \frac{34}{57} = \frac{23}{57}$$

(*e*) $P(\text{1 of each color is drawn}) = \dfrac{({}_8C_1)({}_3C_1)({}_9C_1)}{{}_{20}C_3} = \dfrac{18}{95}$

(*f*) $P(\text{balls drawn in order red, white, blue}) = \dfrac{1}{3!}P(\text{1 of each color is drawn})$

$$= \frac{1}{6}\left(\frac{18}{95}\right) = \frac{3}{95}, \text{ using } (e)$$

Another method.

$$P(R_1 \cap W_2 \cap B_3) = P(R_1)\,P(W_2 \mid R_1)\,P(B_3 \mid R_1 \cap W_2)$$
$$= \left(\frac{8}{20}\right)\left(\frac{3}{19}\right)\left(\frac{9}{18}\right) = \frac{3}{95}$$

1.45. In the game of *poker* five cards are drawn from a pack of 52 well-shuffled cards. Find the probability that (*a*) 4 are aces, (*b*) 4 are aces and 1 is a king, (*c*) 3 are tens and 2 are jacks, (*d*) a nine, ten, jack, queen, king are obtained in any order, (*e*) 3 are of any one suit and 2 are of another, (*f*) at least 1 ace is obtained.

(*a*) $P(\text{4 aces}) = \dfrac{({}_4C_4)({}_{48}C_1)}{{}_{52}C_5} = \dfrac{1}{54{,}145}$

(*b*) $P(\text{4 aces and 1 king}) = \dfrac{({}_4C_4)({}_4C_1)}{{}_{52}C_5} = \dfrac{1}{649{,}740}$

(*c*) $P(\text{3 are tens and 2 are jacks}) = \dfrac{({}_4C_3)({}_4C_2)}{{}_{52}C_5} = \dfrac{1}{108{,}290}$

(*d*) $P(\text{nine, ten, jack, queen, king in any order}) = \dfrac{({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)}{{}_{52}C_5} = \dfrac{64}{162{,}435}$

(e) $P(\text{3 of any one suit, 2 of another}) = \dfrac{(4 \cdot {}_{13}C_3)(3 \cdot {}_{13}C_2)}{{}_{52}C_5} = \dfrac{429}{4165}$

since there are 4 ways of choosing the first suit and 3 ways of choosing the second suit.

(f) $P(\text{no ace}) = \dfrac{{}_{48}C_5}{{}_{52}C_5} = \dfrac{35{,}673}{54{,}145}$. Then $P(\text{at least one ace}) = 1 - \dfrac{35{,}673}{54{,}145} = \dfrac{18{,}472}{54{,}145}$.

1.46. Determine the probability of three 6's in five tosses of a fair die.

Let the tosses of the die be represented by the five spaces — — — — —. In each space we will have the events 6 or not-6 (6′). For example, three 6's and two not-6's can occur as 6 6 6′ 6 6′ or 6 6′ 6 6′ 6, etc.

Now the probability of the outcome 6 6 6′ 6 6′ is

$$P(6\,6\,6'\,6\,6') = P(6)\,P(6)\,P(6')\,P(6)\,P(6') = \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

since we assume independence. Similarly

$$P = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

for all other outcomes in which three 6's and two not-6's occur. But there are ${}_5C_3 = 10$ such outcomes and these are mutually exclusive. Hence the required probability is

$$P(6\,6\,6'\,6\,6' \text{ or } 6\,6'\,6\,6'\,6 \text{ or } \ldots) = {}_5C_3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{5!}{3!\,2!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{125}{3888}$$

In general, if $p = P(A)$ and $q = 1 - p = P(A')$, then by using the same reasoning as given above the probability of getting exactly x A's in n independent trials is

$${}_nC_x p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

1.47. A shelf has 6 mathematics books and 4 physics books. Find the probability that 3 particular mathematics books will be together.

All the books can be arranged among themselves in ${}_{10}P_{10} = 10!$ ways. Let us assume that the 3 particular mathematics books actually are replaced by 1 book. Then we have a total of 8 books which can be arranged among themselves in ${}_8P_8 = 8!$ ways. But the 3 mathematics books themselves can be arranged in ${}_3P_3 = 3!$ ways. The required probability is thus given by

$$\frac{8!\,3!}{10!} = \frac{1}{15}$$

STIRLING'S APPROXIMATION TO $n!$

1.48. Evaluate 50!.

For large n, $n! \sim \sqrt{2\pi n}\, n^n e^{-n}$. Thus

$$50! \sim \sqrt{2\pi(50)}\, 50^{50} e^{-50} \equiv N$$

To evaluate N use logarithms to the base 10. Then

$$\begin{aligned}
\log N &= \log\left(\sqrt{100\pi}\, 50^{50} e^{-50}\right) \\
&= \frac{1}{2}\log 100 + \frac{1}{2}\log \pi + 50 \log 50 - 50 \log e \\
&= \frac{1}{2}\log 100 + \frac{1}{2}\log 3.142 + 50 \log 50 - 50 \log 2.718 \\
&= \frac{1}{2}(2) + \frac{1}{2}(0.4972) + 50(1.6990) - 50(0.4343) = 64.4836
\end{aligned}$$

from which $N = 3.04 \times 10^{64}$, a number which has 65 digits.

MISCELLANEOUS PROBLEMS

1.49. A and B play 12 games of chess of which 6 are won by A, 4 are won by B, and 2 end in a draw. They agree to play a tournament consisting of 3 games. Find the probability that (*a*) A wins all three games, (*b*) two games end in a draw, (*c*) A and B win alternately, (*d*) B wins at least one game.

Let A_1, A_2, A_3 denote the events "A wins" in 1st, 2nd, and 3rd games respectively,

B_1, B_2, B_3 denote the events "B wins" in 1st, 2nd and 3rd games respectively.

On the basis of their past performance (empirical probability) we shall assume that

$$P(\text{A wins any one game}) = \frac{6}{12} = \frac{1}{2} \qquad P(\text{B wins any one game}) = \frac{4}{12} = \frac{1}{3}$$

(*a*) $$P(\text{A wins all 3 games}) = P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2)\,P(A_3) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$$

assuming that the results of each game are independent of the results of any others. (This assumption would not be justifiable if either player were *psychologically influenced* by the other one's winning or losing.)

(*b*) In any one game the probability of a nondraw (i.e. either A or B wins) is $q = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$ and the probability of a draw is $p = 1 - q = \frac{1}{6}$. Then the probability of 2 draws in 3 trials is (see Problem 1.46)

$$\binom{3}{2} p^2 q^{3-2} = 3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{5}{72}$$

(*c*)
$$\begin{aligned} P(\text{A and B win alternately}) &= P(\text{A wins then B wins then A wins} \\ &\qquad \textit{or } \text{B wins then A wins then B wins}) \\ &= P(A_1 \cap B_2 \cap A_3) + P(B_1 \cap A_2 \cap B_3) \\ &= P(A_1)\,P(B_2)\,P(A_3) + P(B_1)\,P(A_2)\,P(B_3) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = \frac{5}{36} \end{aligned}$$

(*d*)
$$\begin{aligned} P(\text{B wins at least one game}) &= 1 - P(\text{B wins no game}) \\ &= 1 - P(B_1' \cap B_2' \cap B_3') \\ &= 1 - P(B_1')\,P(B_2')\,P(B_3') \\ &= 1 - \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{19}{27} \end{aligned}$$

1.50. A and B play a game in which they alternately toss a pair of dice. The one who is first to get a total of 7 wins the game. Find the probability that (*a*) the one who tosses first will win the game, (*b*) the one who tosses second will win the game.

(*a*) The probability of getting a 7 on a single toss of a pair of dice, assumed fair, is $\frac{1}{6}$ as seen from Problem 1.18 and Fig. 1-17. If we suppose that A is the first to toss then A will win in any of the following mutually exclusive cases with indicated associated probabilities:

(1) A wins on 1st toss. Probability = $\frac{1}{6}$.

(2) A loses on 1st toss, B then loses, A then wins. Probability = $(\frac{5}{6})(\frac{5}{6})(\frac{1}{6})$.

(3) A loses on 1st toss, B loses, A loses, B loses, A wins.
Probability = $(\frac{5}{6})(\frac{5}{6})(\frac{5}{6})(\frac{5}{6})(\frac{1}{6})$.

. . . . . .

Then the probability that A wins is

$$\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots$$

$$= \frac{1}{6}\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right] = \frac{1/6}{1 - (5/6)^2} = \frac{6}{11}$$

where we have used the result 6 of Appendix A with $x = (5/6)^2$.

(*b*) The probability that B wins the game is similarly

$$\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots = \left(\frac{5}{6}\right)\left(\frac{1}{6}\right)\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right]$$

$$= \frac{5/36}{1 - (5/6)^2} = \frac{5}{11}$$

Thus we would give 6 to 5 odds that the first one to toss will win. Note that since

$$\frac{6}{11} + \frac{5}{11} = 1$$

the probability of a tie is zero. This would not be true if the game was limited. See Problems 1.151 and 1.152.

1.51. A machine produces a total of 12,000 bolts a day which are on the average 3% defective. Find the probability that out of 600 bolts chosen at random 12 will be defective.

Of the 12,000 bolts, 3% or 360 are defective and 11,640 are not. Then

$$\text{Required probability} = \frac{{}_{360}C_{12}\ {}_{11,640}C_{588}}{{}_{12,000}C_{600}}$$

1.52. A box contains 5 red and 4 white marbles. Two marbles are drawn successively from the box without replacement and it is noted that the second one is white. What is the probability that the first is also white?

Method 1.

If W_1, W_2 are the events "white on 1st draw, "white on 2nd draw," respectively, we are looking for $P(W_1 \mid W_2)$. This is given by

$$P(W_1 \mid W_2) = \frac{P(W_1 \cap W_2)}{P(W_2)} = \frac{(4/9)(3/8)}{4/9} = \frac{3}{8}$$

Method 2.

Since the second is known to be white there are only 3 ways out of the remaining 8 in which the first can be white, so that the probability is 3/8.

1.53. The probabilities that a husband and wife will be alive 20 years from now are given by 0.8 and 0.9 respectively. Find the probability that in 20 years (*a*) both, (*b*) neither, (*c*) at least one, will be alive.

Let H, W be the events that the husband and wife, respectively, will be alive in 20 years. Then $P(H) = 0.8$, $P(W) = 0.9$. We suppose that H and W are independent events, which may or may not be reasonable.

(*a*) $P(\text{both will be alive}) = P(H \cap W) = P(H)\,P(W) = (0.8)(0.9) = 0.72$

(*b*) $P(\text{neither will be alive}) = P(H' \cap W') = P(H')\,P(W') = (0.2)(0.1) = 0.02$

(*c*) $P(\text{at least one will be alive}) = 1 - P(\text{neither will be alive}) = 1 - 0.02 = 0.98$

1.54. An inefficient secretary places n different letters into n differently addressed envelopes at random. Find the probability that at least one of the letters will arrive at the proper destination.

Let $A_1, A_2, \ldots, A_n$ denote the events that the 1st, 2nd, ..., nth letter is in the correct envelope. Then the event that at least one letter is in the correct envelope is $A_1 \cup A_2 \cup \cdots \cup A_n$ and we want to find $P(A_1 \cup A_2 \cup \cdots \cup A_n)$. From a generalization of the results (*10*) and (*11*), page 7, (see Problem 1.79) we have

$$(1) \qquad P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum P(A_k) - \sum P(A_j \cap A_k) + \sum P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

where $\sum P(A_k)$ is the sum of the probabilities of A_k from 1 to n, $\sum P(A_j \cap A_k)$ is the sum of the probabilities of $A_j \cap A_k$ with j and k from 1 to n and $k > j$, etc. We have for example the following:

$$(2) \qquad P(A_1) = \frac{1}{n} \quad \text{and similarly} \quad P(A_k) = \frac{1}{n}$$

since of the n envelopes only 1 will have the proper address. Also

$$(3) \qquad P(A_1 \cap A_2) = P(A_1)\, P(A_2 \mid A_1) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)$$

since if the 1st letter is in the proper envelope then only 1 of the remaining $n-1$ envelopes will be proper. In a similar way we find

$$(4) \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)$$

etc., and finally

$$(5) \qquad P(A_1 \cap A_2 \cap \cdots \cap A_n) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\cdots\left(\frac{1}{1}\right) = \frac{1}{n!}$$

Now in the sum $\sum P(A_j \cap A_k)$ there are $\binom{n}{2} = {}_nC_2$ terms all having the value given by (*3*). Similarly in $\sum P(A_i \cap A_j \cap A_k)$ there are $\binom{n}{3} = {}_nC_3$ terms all having the value given by (*4*). Thus the required probability is

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \binom{n}{1}\left(\frac{1}{n}\right) - \binom{n}{2}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right) + \binom{n}{3}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right) - \cdots + (-1)^{n-1}\binom{n}{n}\left(\frac{1}{n!}\right)$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!}$$

From calculus we know that (see Appendix A)

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

so that for $x = -1$

$$e^{-1} = 1 - \left(1 - \frac{1}{2!} + \frac{1}{3!} - \cdots\right)$$

or

$$1 - \frac{1}{2!} + \frac{1}{3!} - \cdots = 1 - e^{-1}$$

It follows that if n is large the required probability is very nearly $1 - e^{-1} = 0.6321$. This means that there is a good chance of at least one letter arriving at the proper destination. The result is remarkable in that the probability remains practically constant for all $n > 10$. Thus the probability that at least one letter will arrive at its proper destination is practically the same whether n is 10 or 10,000.

1.55. Find the probability that n people ($n \leqq 365$) selected at random will have n different birthdays.

We assume that there are only 365 days in a year and that all birthdays are equally probable, assumptions which are not quite met in reality.

The first of the n people has of course some birthday with probability $365/365 = 1$. Then, if the second is to have a different birthday, it must occur on one of the other 364 days. Thus the probability that the second person has a birthday different from the first is $364/365$. Similarly the probability that the third person has a birthday different from the first two is $363/365$. Finally, the probability that the nth person has a birthday different from the others is $(365 - n + 1)/365$. Thus we have

$$P(\text{all } n \text{ birthdays are different}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365}$$

$$= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right)$$

1.56. Determine how many people are required in Problem 1.55 to make the probability of distinct birthdays less than 1/2.

Denoting the given probability by p and taking natural logarithms we find

$$(1) \qquad \ln p = \ln\left(1 - \frac{1}{365}\right) + \ln\left(1 - \frac{2}{365}\right) + \cdots + \ln\left(1 - \frac{n-1}{365}\right)$$

But we know from calculus (Appendix A, formula 7) that

$$(2) \qquad \ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$$

so that (1) can be written

$$(3) \qquad \ln p = -\left[\frac{1 + 2 + \cdots + (n-1)}{365}\right] - \frac{1}{2}\left[\frac{1^2 + 2^2 + \cdots + (n-1)^2}{(365)^2}\right] - \cdots$$

Using the facts that for $n = 2, 3, \ldots$ (Appendix A, formulas 1 and 2)

$$(4) \qquad 1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}, \qquad 1^2 + 2^2 + \cdots + (n-1)^2 = \frac{n(n-1)(2n-1)}{6}$$

we obtain for (3)

$$(5) \qquad \ln p = -\frac{n(n-1)}{730} - \frac{n(n-1)(2n-1)}{12(365)^2} - \cdots$$

For n small compared to 365, say $n < 30$, the second and higher terms on the right of (5) are negligible compared to the first term, so that a good approximation in this case is

$$(6) \qquad \ln p = -\frac{n(n-1)}{730}$$

For $p = \frac{1}{2}$, $\ln p = -\ln 2 = -0.693$. Thus we have

$$(7) \qquad \frac{n(n-1)}{730} = 0.693 \quad \text{or} \quad n^2 - n - 506 = 0 \quad \text{or} \quad (n - 23)(n + 22) = 0$$

so that $n = 23$. Our conclusion therefore is that if n is larger than 23 we can give better than even odds that at least two people will have the same birthday.

Supplementary Problems

SETS

1.57. Let A be the set of all natural numbers between 5 and 15 which are even. Describe A according to (a) the roster method, (b) the property method.

1.58. Let $A = \{x \mid x^2 - 3x + 2 = 0\}$, $B = \{x \mid x^2 \leqq 16\}$. Determine whether or not $A \subset B$.

1.59. Prove that for any set A we have $A \subset A$.

1.60. Discuss the truth or falsity of the following statements. (a) If A and B are any two sets then either $A \subset B$, $A \supset B$, or $A = B$. (b) If x and y are any two real numbers than either $x < y$, $x > y$, or $x = y$.

1.61. Prove that any subset of the empty set must be the empty set.

1.62. Let $\mathcal{S}$ be the set or class of all sets which are not elements of themselves. (*a*) Prove that if $\mathcal{S} \in \mathcal{S}$, then $\mathcal{S} \notin \mathcal{S}$. (*b*) Prove that if $\mathcal{S} \notin \mathcal{S}$, then $\mathcal{S} \in \mathcal{S}$. The paradox described is called *Russell's paradox.*

SET OPERATIONS, VENN DIAGRAMS AND THEOREMS ON SETS

1.63. Let a universe be given by $\mathcal{U} = \{1, 2, 3, 4, 5\}$ and suppose that subsets of $\mathcal{U}$ are $A = \{1, 5\}$, $B = \{2, 5, 3\}$, $C = \{4, 2\}$. Find (*a*) $A \cup (B \cup C)$, (*b*) $(A \cup B) \cup C$, (*c*) $A \cap (B \cup C)$, (*d*) $(A \cap B) \cup (A \cap C)$, (*e*) $A' \cap (B' \cap C')$, (*f*) $(A \cup B) - (A \cup C)$, (*g*) $(A \cap C)' \cup B$, (*h*) $A - (B' \cup C')$.

1.64. Let $\mathcal{U}$ be the set of all nonnegative integers and consider the subsets

$$A = \{x \mid x \text{ is an even integer, } 1 \leq x < 6\} \quad \text{and} \quad B = \{x \mid x \text{ is a prime number, } 0 < x \leq 4\}$$

Find (*a*) $A \cup B$, (*b*) $A \cap B$, (*c*) $A' \cap B'$, (*d*) $A - B$, (*e*) $B - A$, (*f*) $(A - B) \cup (B - A)$.

1.65. Use a Venn diagram to sketch each of the following sets:

(*a*) $A \cap (B \cup C)$ (*c*) $A' \cap (B \cup C)'$ (*e*) $A' - (B \cup C)'$

(*b*) $A \cup (B \cap C)$ (*d*) $A - (B \cap C)$

1.66. Is $(A - B)' = A' - B'$? Justify your answer.

1.67. Prove (*a*) Theorem 1-2, page 3, (*b*) Theorem 1-3, page 3, (*c*) Theorem 1-4, page 3.

1.68. (*a*) Prove De Morgan's second law, Theorem 1-12*b*, page 3, and (*b*) illustrate by using a Venn diagram.

1.69. Generalize De Morgan's first and second laws to any number of sets. (See Problem 1.7.)

1.70. Illustrate the principle of duality by referring to the theorems on page 3.

1.71. Prove that $(A - B) \cup B = A$ if and only if $B \subset A$, and illustrate by using a Venn diagram.

1.72. Prove or disprove: If $A - B = \emptyset$, then $A = B$.

1.73. Prove that $A \cup B = [A - (A \cap B)] \cup [B - (A \cap B)] \cup (A \cap B)$ and illustrate by a Venn diagram.

1.74. Generalize the result of Problem 1.9.

RANDOM EXPERIMENTS, SAMPLE SPACES AND EVENTS

1.75. Describe a sample space for each of the following random experiments: (*a*) 3 tosses of a coin, (*b*) the number of smokers in a group of 500 adult males, (*c*) tossing a coin until a tail appears, (*d*) the number of incoming calls on a telephone trunkline, (*e*) the number of nuclear particles entering a Geiger counter, (*f*) tossing a coin and a die.

1.76. An experiment consists of tossing a coin and a die. If A is the event "heads comes up" in tossing the coin and B is the event "3 or 6 comes up" in tossing the die, state in words the meaning of each of the following: (*a*) A', (*b*) B', (*c*) $A \cap B$, (*d*) $A \cap B'$, (*e*) $A - B$, (*f*) $B - A$, (*g*) $A' \cup B$.

THEOREMS ON PROBABILITY

1.77. Complete the proof in Problem 1.14(*b*) by proving (without a Venn diagram) that

$$A \cup B = A \cup [B - (A \cap B)]$$

where A and $B - (A \cap B)$ are mutually exclusive.

1.78. Prove the result (*11*), page 7.

1.79. Generalize the results (*10*) and (*11*), page 7, and thus prove the result (*1*) of Problem 1.54, page 30.

1.80. Prove that $P(A' \cup B') = 1 - P(A \cap B)$.

CALCULATION OF PROBABILITIES

1.81. Determine the probability p, or an estimate of it, for each of the following events:

(*a*) A king, ace, jack of clubs or queen of diamonds appears in drawing a single card from a well-shuffled ordinary deck of cards.

(*b*) The sum 8 appears in a single toss of a pair of fair dice.

(*c*) A nondefective bolt will be found next if out of 600 bolts already examined, 12 were defective.

(*d*) A 7 or 11 comes up in a single toss of a pair of fair dice.

(*e*) At least one head appears in three tosses of a fair coin.

1.82. An experiment consists of drawing three cards in succession from a well-shuffled ordinary deck of cards. Let A_1 be the event "king on first draw," A_2 the event "king on second draw," and A_3 the event "king on third draw." State in words the meaning of each of the following:
(*a*) $P(A_1 \cap A_2')$, (*b*) $P(A_1 \cup A_2)$, (*c*) $P(A_1' \cup A_2')$, (*d*) $P(A_1' \cap A_2' \cap A_3')$, (*e*) $P[(A_1 \cap A_2) \cup (A_2' \cap A_3)]$.

1.83. A marble is drawn at random from a box containing 10 red, 30 white, 20 blue and 15 orange marbles. Find the probability that it is (*a*) orange or red, (*b*) not red or blue, (*c*) not blue, (*d*) white, (*e*) red, white or blue.

1.84. Two marbles are drawn in succession from the box of Problem 1.83, replacement being made after each drawing. Find the probability that (*a*) both are white, (*b*) the first is red and the second is white, (*c*) neither is orange, (*d*) they are either red or white or both (red and white), (*e*) the second is not blue, (*f*) the first is orange, (*g*) at least one is blue, (*h*) at most one is red, (*i*) the first is white but the second is not, (*j*) only one is red.

1.85. Work Problem 1.84 if there is no replacement after each drawing.

CONDITIONAL PROBABILITY AND INDEPENDENT EVENTS

1.86. A box contains 2 red and 3 blue marbles. Find the probability that if two marbles are drawn at random (without replacement) (*a*) both are blue, (*b*) both are red, (*c*) one is red and one is blue.

1.87. Find the probability of drawing 3 aces at random from a deck of 52 ordinary cards if the cards are (*a*) replaced, (*b*) not replaced.

1.88. If at least one child in a family with two children is a boy what is the probability that both children are boys?

1.89. Show that conditional probability as defined by (*17*), page 8, satisfies the axioms of probability on page 6 and thus all the theorems of probability.

1.90. Prove that if $P(A) > P(B)$ then $P(A \mid B) > P(B \mid A)$.

1.91. If A is independent of B prove that (*a*) A is independent of B', (*b*) A' is independent of B'.

1.92. If A, B, C are independent events prove that (*a*) A and $B \cup C$, (*b*) A and $B \cap C$, (*c*) A and $B - C$, are independent.

1.93. Let A_1 = event "odd number on first die," A_2 = event "odd number on second die," A_3 = event "odd total on both dice." Show that A_1, A_2; A_2, A_3; A_1, A_3 are independent but that A_1, A_2, A_3 are not independent.

1.94. Box *I* contains 3 red and 5 white balls, while Box *II* contains 4 red and 2 white balls. A ball is chosen at random from the first box and placed in the second box without observing its color. Then a ball is drawn from the second box. Find the probability that it is white.

BAYES' THEOREM OR RULE

1.95. A box contains 3 blue and 2 red marbles while another box contains 2 blue and 5 red marbles. A marble drawn at random from one of the boxes turns out to be blue. What is the probability that it came from the first box?

1.96. Each of three identical jewelry boxes has two drawers. In each drawer of the first box there is a gold watch. In each drawer of the second box there is a silver watch. In one drawer of the third box there is a gold watch while in the other there is a silver watch. If we select a box at random, open one of the drawers and find it to contain a silver watch, what is the probability that the other drawer has the gold watch?

1.97. Urn I has 2 white and 3 black balls; Urn II, 4 white and 1 black; and Urn III, 3 white and 4 black. An urn is selected at random and a ball drawn at random is found to be white. Find the probability that Urn I was selected.

COMBINATORIAL ANALYSIS, COUNTING AND TREE DIAGRAMS

1.98. A coin is tossed three times. Use a tree diagram to determine the various possibilities which can arise.

1.99. Three cards are drawn at random (without replacement) from an ordinary deck of 52 cards. Use a tree diagram to determine the number of ways in which one can draw (*a*) a diamond and a club and a heart in succession, (*b*) two hearts and then a club or a spade.

1.100. In how many ways can 3 different coins be placed in 2 different purses?

PERMUTATIONS

1.101. Evaluate (*a*) ${}_4P_2$, (*b*) ${}_7P_5$, (*c*) ${}_{10}P_3$.

1.102. For what value of n is ${}_{n+1}P_3 = {}_nP_4$?

1.103. In how many ways can 5 people be seated on a sofa if there are only 3 seats available?

1.104. In how many ways can 7 books be arranged on a shelf if (*a*) any arrangement is possible, (*b*) 3 particular books must always stand together, (*c*) two particular books must occupy the ends?

1.105. How many numbers consisting of five different digits each can be made from the digits $1, 2, 3, \ldots, 9$ if (*a*) the numbers must be odd, (*b*) the first two digits of each number are even?

1.106. Solve Problem 1.105 if repetitions of the digits are allowed.

1.107. How many different three-digit numbers can be made with 3 fours, 4 twos and 2 threes?

1.108. In how many ways can 3 men and 3 women be seated at a round table if (*a*) no restriction is imposed, (*b*) two particular women must not sit together, (*c*) each woman is to be between two men?

COMBINATIONS

1.109. Evaluate (*a*) ${}_5C_3$, (*b*) ${}_8C_4$, (*c*) ${}_{10}C_8$.

1.110. For what value of n is $3 \cdot {}_{n+1}C_3 = 7 \cdot {}_nC_2$?

1.111. In how many ways can 6 questions be selected out of 10?

1.112. How many different committees of 3 men and 4 women can be formed from 8 men and 6 women?

1.113. In how many ways can 2 men, 4 women, 3 boys and 3 girls be selected from 6 men, 8 women, 4 boys and 5 girls if (*a*) no restrictions are imposed, (*b*) a particular man and woman must be selected?

1.114. In how many ways can a group of 10 people be divided into (*a*) two groups consisting of 7 and 3 people, (*b*) three groups consisting of 5, 3 and 2 people?

1.115. From 5 statisticians and 6 economists a committee consisting of 3 statisticians and 2 economists is to be formed. How many different committees can be formed if (*a*) no restrictions are imposed, (*b*) two particular statisticians must be on the committee, (*c*) one particular economist cannot be on the committee?

1.116. Find the number of (*a*) combinations and (*b*) permutations of four letters each that can be made from the letters of the word *Tennessee*.

BINOMIAL COEFFICIENTS

1.117. Calculate (*a*) ${}_6C_3$, (*b*) $\binom{11}{4}$, (*c*) $({}_8C_2)({}_4C_3)/{}_{12}C_5$.

1.118. Expand (*a*) $(x+y)^6$, (*b*) $(x-y)^4$, (*c*) $(x-x^{-1})^5$, (*d*) $(x^2+2)^4$.

1.119. Find the coefficient of x in $\left(x+\frac{2}{x}\right)^9$.

1.120. Show that (*a*) $\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n} = 2^n$

(*b*) $\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n} = 0$

1.121. Prove that (*a*) $\sum_{j=1}^{n} j({}_nC_j) = n\cdot 2^{n-1}$, (*b*) $\sum_{j=1}^{n} (-1)^{j-1} j({}_nC_j) = 0$.

PROBABILITY USING COMBINATORIAL ANALYSIS

1.122. Find the probability of scoring a total of 7 points (*a*) once, (*b*) at least once, (*c*) twice, in two tosses of a pair of fair dice.

1.123. Two cards are drawn successively from an ordinary deck of 52 well-shuffled cards. Find the probability that (*a*) the first card is not a ten of clubs or an ace, (*b*) the first card is an ace but the second is not, (*c*) at least one card is a diamond, (*d*) the cards are not of the same suit, (*e*) not more than one card is a picture card (jack, queen, king), (*f*) the second card is not a picture card, (*g*) the second card is not a picture card given that the first was a picture card, (*h*) the cards are picture cards or spades or both.

1.124. A box contains 9 tickets numbered from 1 to 9 inclusive. If 3 tickets are drawn from the box one at a time, find the probability that they are alternately either odd, even, odd or even, odd, even.

1.125. The odds in favor of A winning a game of chess against B are 3 : 2. If three games are to be played, what are the odds (*a*) in favor of A winning at least two games out of the three, (*b*) against A losing the first two games to B?

1.126. In the game of *bridge* each of 4 players is dealt 13 cards from an ordinary well-shuffled deck of 52 cards. Find the probability that one of the players (say, the eldest) gets (*a*) 7 diamonds, 2 clubs, 3 hearts and 1 spade; (*b*) a complete suit.

1.127. An urn contains 6 red and 8 blue marbles. Five marbles are drawn at random from it without replacement. Find the probability that 3 are red and 2 are blue.

1.128. (*a*) Find the probability of getting the sum 7 on at least one of three tosses of a pair of fair dice. (*b*) How many tosses are needed in order that the probability in (*a*) be greater than 0.95?

1.129 Three cards are drawn from an ordinary deck of 52 cards. Find the probability that (*a*) all cards are of one suit, (*b*) at least two aces are drawn.

1.130. Find the probability that a bridge player is given 13 cards of which 9 cards are of one suit.

STIRLING'S APPROXIMATION TO $n!$

1.131. In how many ways can 30 individuals be selected out of 100?

1.132. Show that ${}_{2n}C_n = 2^{2n}/\sqrt{\pi n}$ approximately, for large values of n.

1.133. Find the percentage error in Stirling's formula for $n = 10$.

1.134. Obtain an approximation to the result of Problem 1.51.

MISCELLANEOUS PROBLEMS

1.135. A sample space consists of 3 sample points with associated probabilities given by $2p$, p^2 and $4p - 1$. Find the value of p.

1.136. Prove that if $A \subset B'$ then $A \cap B = \emptyset$.

1.137. Prove that $A - (A \cap B) = A \cap B'$.

1.138. How many words can be made from 5 letters if (*a*) all letters are different, (*b*) 2 letters are identical, (*c*) all letters are different but 2 particular letters cannot be adjacent?

1.139. Four integers are chosen at random between 0 and 9 inclusive. Find the probability that (*a*) they are all different, (*b*) not more than two are the same.

1.140. A pair of dice is tossed repeatedly. Find the probability that an 11 occurs for the first time on the 6th toss.

1.141. What is the least number of tosses needed in Problem 1.140 so that the probability of getting an 11 will be greater than (*a*) 0.5, (*b*) 0.95?

1.142. Discuss the following: There is no such thing as a fair coin since in any number of tosses it is extremely unlikely that the number of heads and tails will be the same.

1.143. Suppose that in tossing a coin 500 times there is a sequence of 24 tosses which result in "heads." Can the coin be fair? Explain.

1.144. Prove that for any events $A_1, A_2, \ldots, A_n$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n)$$

1.145. In tossing a pair of dice the sum can be 2, 3, ..., 12. Could we assign probabilities of 1/11 to each of these? Explain.

1.146. In a game of poker find the probability of getting (*a*) a *royal flush*, which consists of the ten, jack, queen, king and ace of a single suit; (*b*) a *full house*, which consists of 3 cards of one face value and 2 of another (such as 3 tens and 2 jacks, etc.); (*c*) all different cards; (*d*) 4 aces.

1.147. The probability that a man will hit a target is $\frac{2}{3}$. If he shoots at the target until he hits it for the first time, find the probability that it will take 5 shots.

1.148. (*a*) A shelf contains 6 separate compartments. In how many ways can 4 indistinguishable marbles be placed in the compartments? (*b*) Work the problem if there are n compartments and r marbles. This type of problem arises in physics in connection with *Bose-Einstein statistics*.

1.149. (*a*) A shelf contains 6 separate compartments. In how many ways can 12 indistinguishable marbles be placed in the compartments so that no compartment is empty? (*b*) Work the problem if there are n compartments and r marbles where $r > n$. This type of problem arises in physics in connection with *Fermi-Dirac statistics*.

1.150. A poker player has cards 2, 3, 4, 6, 8. He wishes to discard the 8 and replace it by another card which he hopes will be a 5 (in which case he gets an "inside straight"). What is the probability that he will succeed assuming that the other three players together have (*a*) one 5, (*b*) two 5's, (*c*) three 5's, (*d*) no 5? Can the problem be worked if the number of 5's in the other players' hands is unknown? Explain.

1.151. Work Problem 1.50 if the game is limited to 3 tosses.

1.152. Generalize the result of Problem 1.151.

1.153. Find the probability that in a game of bridge (*a*) two, (*b*) three, (*c*) all four players have a complete suit.

1.154. Prove that $\binom{n}{r} = \sum_{j=0}^{r} \binom{k}{j}\binom{n-k}{r-j}$ and give a combinatorial interpretation. (*Hint*: Consider $(1+x)^k(1+x)^{n-k}$ and find the coefficient of x^j in the product.)

1.155. Prove that $\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2$ and give a combinatorial interpretation.

1.156. Prove that the probability that the secretary of Problem 1.54 will get exactly a letters in the correct envelopes is $\frac{1}{a!}\sum_{k=0}^{n-a}\frac{(-1)^k}{k!}$. [*Hint*: Denoting the desired probability as $p_n(a)$, show that $p_n(a) = \frac{1}{a!}p_{n-a}(0)$ and then use the result of Problem 1.54.]

Ziyad Asaad

Chapter 2

Random Variables and Probability Distributions

RANDOM VARIABLES

Suppose that to each point of a sample space we assign a number. We then have a *function* defined on the sample space. This function is called a *random variable* (or *stochastic variable*) or more precisely a *random function* (*stochastic function*). It is usually denoted by a capital letter such as X or Y. In general a random variable has some specified physical, geometrical or other significance.

EXAMPLE 2.1. Suppose that a coin is tossed twice so that the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads which can come up. With each sample point we can associate a number for X as shown in Table 2-1. Thus for example in the case of HH (i.e. 2 heads) $X = 2$ while for TH (1 head) $X = 1$. It follows that X is a random variable.

Table 2-1

Sample Point	HH	HT	TH	TT
X	2	1	1	0

It should be noted that many other random variables could also be defined on this sample space, for example the square of the number of heads, the number of heads minus the number of tails, etc.

A random variable which takes on a finite or countably infinite number of values (see page 4) is called a *discrete random variable* while one which takes on a noncountably infinite number of values is called a *nondiscrete* or *continuous random variable.*

DISCRETE PROBABILITY DISTRIBUTIONS

Let X be a discrete random variable and suppose that the possible values which it can assume are given by $x_1, x_2, x_3, \ldots$, arranged in increasing order of magnitude. Suppose also that these values are assumed with probabilities given by

$$P(X = x_k) = f(x_k) \qquad k = 1, 2, \ldots \tag{1}$$

It is convenient to introduce the *probability function*, also referred to as *probability distribution*, given by

$$P(X = x) = f(x) \tag{2}$$

For $x = x_k$ this reduces to (1) while for other values of x, $f(x) = 0$.

In general $f(x)$ is a probability function if

1. $f(x) \geqq 0$
2. $\sum_x f(x) = 1$

where the sum in 2 is taken over all possible values of x. A graph of $f(x)$ is called a *probability graph.*

EXAMPLE 2.2. (*a*) Find the probability function corresponding to the random variable X of Example 2.1 and (*b*) construct a probability graph.

(*a*) Assuming that the coin is fair we have

$$P(HH) = \frac{1}{4} \qquad P(HT) = \frac{1}{4} \qquad P(TH) = \frac{1}{4} \qquad P(TT) = \frac{1}{4}$$

Then

$$P(X=0) = P(TT) = \frac{1}{4}$$

$$P(X=1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X=2) = P(HH) = \frac{1}{4}$$

The probability function is thus given by Table 2-2.

Table 2-2

x	0	1	2
$f(x)$	1/4	1/2	1/4

(*b*) A probability graph can be obtained by use of a *bar chart*, as indicated in Fig. 2-1, or a *histogram*, as indicated in Fig. 2-2. In the bar chart the sum of the ordinates is 1 while in the histogram the sum of the rectangular areas is 1. In the case of the histogram we can think of the random variable X as being made continuous, e.g. $X = 1$ means that it lies between 0.5 and 1.5.

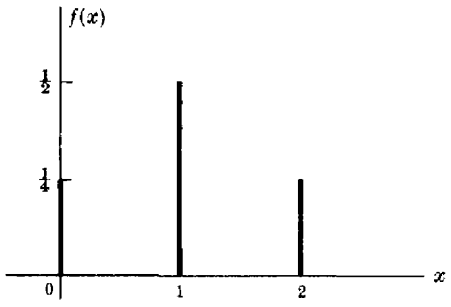


Fig. 2-1. Bar Chart

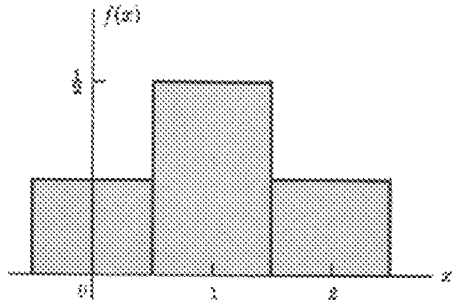


Fig. 2-2. Histogram

DISTRIBUTION FUNCTIONS FOR DISCRETE RANDOM VARIABLES

The *cumulative distribution function*, or briefly the *distribution function*, for a random variable X is defined by

$$P(X \leqq x) = F(x) \tag{3}$$

where x is any real number, i.e. $-\infty < x < \infty$. The distribution function can be obtained from the probability function by noting that

$$F(x) = P(X \leqq x) = \sum_{u \leqq x} f(u) \tag{4}$$

where the sum on the right is taken over all values of u for which $u \leqq x$. Conversely the probability function can be obtained from the distribution function.

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$ then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leqq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leqq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leqq x < \infty \end{cases} \tag{5}$$

EXAMPLE 2.3. (*a*) Find the distribution function for the random variable X of Example 2.2. (*b*) Obtain its graph.

(*a*) The distribution function is

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

(*b*) The graph of $F(x)$ is shown in Fig. 2-3.

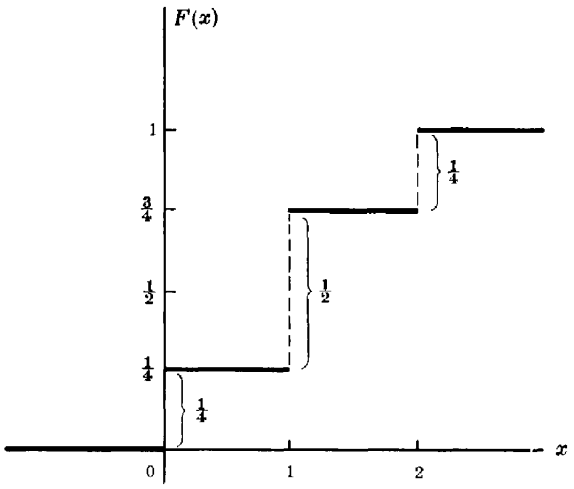


Fig. 2-3

The following things about the above distribution function, which are true in general, should be noted.

1. The magnitudes of the jumps at 0, 1, 2 are $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ which are precisely the ordinates in Fig. 2-1. This fact enables one to obtain the probability function from the distribution function.
2. Because of the appearance of the graph of Fig. 2-3 it is often called a *staircase function* or *step function.* The value of the function at an integer is obtained from the higher step, thus the value at 1 is $\frac{3}{4}$ and not $\frac{1}{4}$. This is expressed mathematically by stating that the distribution function is *continuous from the right* at 0, 1, 2.
3. As we proceed from left to right (i.e. going *upstairs*) the distribution function either remains the same or increases, taking on values from 0 to 1. Because of this it is said to be a *monotonically increasing function.*

CONTINUOUS PROBABILITY DISTRIBUTIONS

If X is a continuous random variable the probability that X takes on any one particular value is generally zero. Therefore we cannot define a probability function in the same way as for a discrete random variable (page 38). In order to arrive at a probability distribution for a continuous random variable we note that the probability that X lies *between two different values* is meaningful.

EXAMPLE 2.4. If an individual is selected at random from a large group of adult males, the probability that his height X is precisely 68 inches (i.e. 68.000... inches) would be zero. However there is a probability greater than zero that X is between 67.500... inches and 68.500... inches, for example.

These ideas and the analogy of Properties 1 and 2, page 38, lead us to postulate the existence of a function $f(x)$ such that

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

where the second is a mathematical statement of the fact that a real-valued random variable must certainly lie between $-\infty$ and ∞. We then define the probability that X lies between a and b by

$$P(a < X < b) = \int_a^b f(x)\,dx \tag{6}$$

We can show that this definition satisfies the axioms of probability given on page 6.

A function $f(x)$ which satisfies the above requirements is called a *probability function* or *probability distribution* for a continuous random variable, but it is more often called a *probability density function* or simply *density function.* Any function $f(x)$ satisfying Properties 1 and 2 above will automatically be a density function and required probabilities can then be obtained from (6).

EXAMPLE 2.5. (a) Find the constant c such that the function

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

is a density function and (b) compute $P(1 < X < 2)$.

(a) Since $f(x)$ satisfies Property 1 if $c \geq 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^3 cx^2\,dx = \frac{cx^3}{3}\Big|_0^3 = 9c$$

and since this must equal 1 we have $c = 1/9$.

(b) $$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \frac{x^3}{27}\Big|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

In case $f(x)$ is continuous, which we shall assume unless otherwise stated, the probability that X is equal to any particular value is zero. In such case we can replace either or both of the signs $<$ in (6) by $\leq$. Thus, in Example 2.5,

$$P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2) = \frac{7}{27}$$

DISTRIBUTION FUNCTIONS FOR CONTINUOUS RANDOM VARIABLES

By analogy with (4), page 39, we define the *distribution function $F(x)$ for a continuous random variable* by

$$F(x) = P(X \leq x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(u)\,du \tag{7}$$

At points of continuity of $f(x)$, the $\leq$ in (7) can be replaced by $<$ if desired.

EXAMPLE 2.6. (a) Find the distribution function for the random variable of Example 2.5. (b) Use the result of (a) to find $P(1 < x \leq 2)$.

(a) We have

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

If $x < 0$ then $F(x) = 0$. If $0 \leq x < 3$ then

$$F(x) = \int_0^x f(u)\,du = \int_0^x \frac{1}{9}u^2\,du = \frac{x^3}{27}$$

If $x \geq 3$ then

$$F(x) = \int_0^3 f(u)\,du + \int_3^x f(u)\,du = \int_0^3 \frac{1}{9}u^2\,du + \int_3^x 0\,du = 1$$

Thus the required distribution function is

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Note that $F(x)$ increases monotonically from 0 to 1 as is required for a distribution function. It should also be noted that $F(x)$ in this case is continuous.

(*b*) We have

$$\begin{aligned} P(1 < X \leqq 2) &= P(X \leqq 2) - P(X \leqq 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

as in Example 2.5.

The probability that X is between x and $x + \Delta x$ is given by

$$P(x \leqq X \leqq x + \Delta x) = \int_x^{x+\Delta x} f(u)\,du \tag{8}$$

so that if Δx is small we have approximately

$$P(x \leqq X \leqq x + \Delta x) = f(x)\Delta x \tag{9}$$

We also see from (7) on differentiating both sides that

$$\frac{dF(x)}{dx} = f(x) \tag{10}$$

at all points where $f(x)$ is continuous, i.e. the derivative of the distribution function is the density function.

LEIBNIZ'S RULE

In order to obtain (*10*) we have used the fact familiar from calculus that

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x) \tag{11}$$

This is a special case of *Leibniz's rule for differentiation of an integral*:

$$\frac{d}{dx}\int_{a_1(x)}^{a_2(x)} F(u, x)\,du = \int_{a_1(x)}^{a_2(x)} \frac{\partial F}{\partial x}du + F(a_2(x), x)\frac{da_2}{dx} - F(a_1(x), x)\frac{da_1}{dx} \tag{12}$$

where a_1, a_2 and F are assumed to be differentiable with respect to x.

GRAPHICAL INTERPRETATIONS

If $f(x)$ is the density function for a random variable X then we can represent $y = f(x)$ graphically by a curve as in Fig. 2-4. Since $f(x) \geqq 0$ the curve cannot fall below the x-axis. The entire area bounded by the curve and the x-axis must be 1 because of Property 2 on page 40. Geometrically the probability that X is between a and b, i.e. $P(a < X < b)$, is then represented by the area shown shaded in Fig. 2-4.

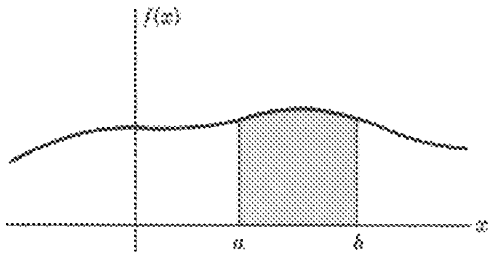


Fig. 2-4

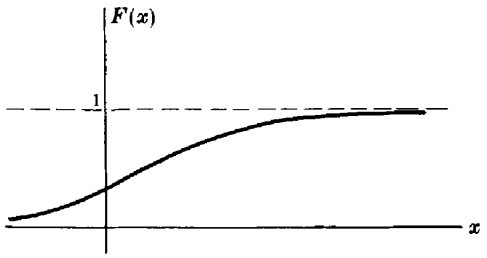


Fig. 2-5

The distribution function $F(x) = P(X \leqq x)$ is a monotonically increasing function which increases from 0 to 1 and is represented by a curve as in Fig. 2-5.

JOINT DISTRIBUTIONS

The above ideas are easily generalized to two or more random variables. We consider the typical case of two random variables which are either both discrete or both continuous. In cases where one variable is discrete and the other continuous, appropriate modifications are easily made. Generalizations to more than two variables can also be made.

1. Discrete case.

If X and Y are two discrete random variables we define the *joint probability function* of X and Y by

$$P(X = x, Y = y) = f(x, y) \tag{13}$$

where 1. $f(x, y) \geqq 0$

2. $\sum_x \sum_y f(x, y) = 1$

i.e. the sum over all values of x and y is one.

Suppose that X can assume any one of m values $x_1, x_2, \ldots, x_m$ and Y can assume any one of n values $y_1, y_2, \ldots, y_n$. Then the probability of the event that $X = x_j$ and $Y = y_k$ is given by

$$P(X = x_j, Y = y_k) = f(x_j, y_k) \tag{14}$$

A joint probability function for X and Y can be represented by a *joint probability table* as in Table 2-3. The probability that $X = x_j$ is obtained by adding all entries in the row corresponding to x_j and is given by

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^{n} f(x_j, y_k) \tag{15}$$

For $j = 1, 2, \ldots, m$ these are indicated by the entry totals in the extreme right-hand column or margin of Table 2-3. Similarly the probability that $Y = y_k$ is obtained by adding all entries in the column corresponding to y_k and is given by

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k) \tag{16}$$

For $k = 1, 2, \ldots, n$ these are indicated by the entry totals in the bottom row or margin of Table 2-3.

Table 2-3

X \ Y	y_1	y_2	$\cdots$	y_n	Totals ↓
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\cdots$	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\cdots$	$f(x_2, y_n)$	$f_1(x_2)$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\cdots$	$f(x_m, y_n)$	$f_1(x_m)$
Totals →	$f_2(y_1)$	$f_2(y_2)$	$\cdots$	$f_2(y_n)$	1 ← Grand Total

Because the probabilities (*15*) and (*16*) are obtained from the margins of the table we often refer to $f_1(x_j)$ and $f_2(y_k)$ (or simply $f_1(x)$ and $f_2(y)$) as the *marginal probability functions* of X and Y respectively. It should also be noted that

$$\sum_{j=1}^{m} f_1(x_j) = 1 \qquad \sum_{k=1}^{n} f_2(y_k) = 1 \tag{17}$$

which can be written

$$\sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j, y_k) = 1 \tag{18}$$

This is simply the statement that the total probability of all entries is 1. The *grand total* of 1 is indicated in the lower right-hand corner of the table.

The *joint distribution function* of X and Y is defined by

$$F(x, y) = P(X \leqq x, Y \leqq y) = \sum_{u \leqq x}\sum_{v \leqq y} f(u, v) \tag{19}$$

In Table 2-3, $F(x, y)$ is the sum of all entries for which $x_j \leqq x$ and $y_k \leqq y$.

2. Continuous case.

The case where both variables are continuous is obtained easily by analogy with the discrete case on replacing sums by integrals. Thus the *joint probability function* for the random variables X and Y (or, as it is more commonly called, the *joint density function* of X and Y) is defined by

1. $f(x, y) \geqq 0$
2. $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$

Graphically $z = f(x, y)$ represents a surface, called the *probability surface*, as indicated in Fig. 2-6. The total volume bounded by this surface and the xy-plane is equal to 1 in accordance with Property 2 above. The probability that X lies between a and b while Y lies between c and d is given graphically by the shaded volume of Fig. 2-6 and mathematically by

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b}\int_{y=c}^{d} f(x, y)\, dx\, dy \tag{20}$$

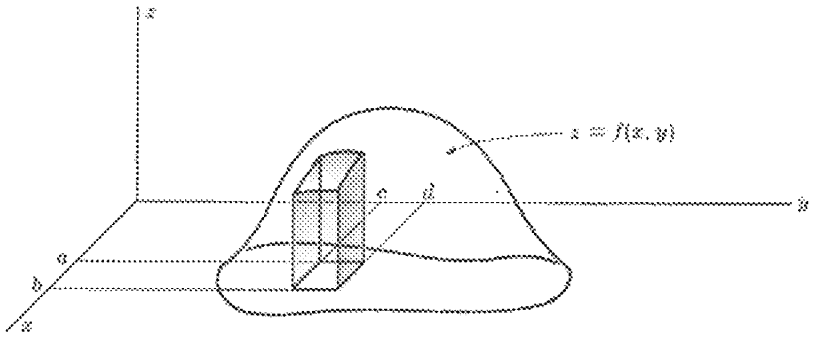


Fig. 2-6

More generally if A represents any event there will be a region $\mathcal{R}_A$ of the xy-plane which corresponds to it. In such case we can find the probability of A by performing the integration over $\mathcal{R}_A$, i.e.

$$P(A) = \iint_{\mathcal{R}_A} f(x, y)\, dx\, dy \tag{21}$$

The *joint distribution function* of X and Y in this case is defined by

$$F(x, y) = P(X \leqq x, Y \leqq y) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{y} f(u, v)\, du\, dv \tag{22}$$

It follows in analogy with (*10*), page 42, that

$$\frac{\partial^2 F}{\partial x\, \partial y} = f(x, y) \tag{23}$$

i.e. the density function is obtained by differentiating the distribution function with respect to x and y.

From (*22*) we obtain

$$P(X \leqq x) = F_1(x) = \int_{u=-\infty}^{x} \int_{v=-\infty}^{\infty} f(u, v)\, du\, dv \tag{24}$$

$$P(Y \leqq y) = F_2(y) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{y} f(u, v)\, du\, dv \tag{25}$$

We call (*24*) and (*25*) the *marginal distribution functions*, or simply the *distribution functions*, of X and Y respectively. The derivatives of (*24*) and (*25*) with respect to x and y are then called the *marginal density functions*, or simply the *density functions*, of X and Y and are given by

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x, v)\, dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u, y)\, du \tag{26}$$

INDEPENDENT RANDOM VARIABLES

Suppose that X and Y are discrete random variables. If the events $X = x$ and $Y = y$ are independent events for all x and y, then we say that X and Y are *independent random variables*. In such case

$$P(X = x, Y = y) = P(X = x)\, P(Y = y) \tag{27}$$

or equivalently

$$f(x, y) = f_1(x)\, f_2(y) \tag{28}$$

Conversely, if for all x and y the joint probability function $f(x, y)$ can be expressed as the product of a function of x alone and a function of y alone (which are then the marginal probability functions of X and Y), X and Y are independent. If however $f(x, y)$ cannot be so expressed then X and Y are *dependent*.

If X and Y are continuous random variables we say that they are *independent random variables* if the events $X \leqq x$ and $Y \leqq y$ are independent events for all x and y. In such case we can write

$$P(X \leqq x, Y \leqq y) = P(X \leqq x)\, P(Y \leqq y) \tag{29}$$

or equivalently

$$F(x, y) = F_1(x)\, F_2(y) \tag{30}$$

where $F_1(x)$ and $F_2(y)$ are the (marginal) distribution functions of X and Y respectively. Conversely, X and Y are independent random variables if for all x and y their joint dis-

tribution function $F(x, y)$ can be expressed as a product of a function of x alone and a function of y alone (which are the marginal distributions of X and Y respectively). If however $F(x, y)$ cannot be so expressed then X and Y are dependent.

For continuous independent random variables it is also true that the joint density function $f(x, y)$ is the product of a function of x alone, $f_1(x)$, and a function of y alone, $f_2(y)$, and these are the (marginal) density functions of X and Y respectively.

CHANGE OF VARIABLES

Given the probability distributions of one or more random variables we are often interested in finding distributions of other random variables which depend on them in some specified manner. Procedures for obtaining these distributions are presented in the following theorems for the case of discrete and continuous variables.

1. Discrete variables.

Theorem 2-1: Let X be a discrete random variable whose probability function is $f(x)$. Suppose that a discrete random variable U is defined in terms of X by $U = \phi(X)$, where to each value of X there corresponds one and only one value of U and conversely, so that $X = \psi(U)$. Then the probability function for U is given by

$$g(u) = f[\psi(u)] \tag{31}$$

Theorem 2-2: Let X and Y be discrete random variables having joint probability function $f(x, y)$. Suppose that two discrete random variables U and V are defined in terms of X and Y by $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$, where to each pair of values of X and Y there corresponds one and only one pair of values of U and V and conversely, so that $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$. Then the joint probability function of U and V is given by

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)] \tag{32}$$

2. Continuous variables.

Theorem 2-3: Let X be a continuous random variable with probability density $f(x)$. Let us define $U = \phi(X)$ where $X = \psi(U)$ as in Theorem 2-1. Then the probability density of U is given by $g(u)$ where

$$g(u)\,|du| = f(x)\,|dx| \tag{33}$$

or

$$g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(u)]\,|\psi'(u)| \tag{34}$$

Theorem 2-4: Let X and Y be continuous random variables having joint density function $f(x, y)$. Let us define $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$ where $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$ as in Theorem 2-2. Then the joint density function of U and V is given by $g(u, v)$ where

$$g(u, v)\,|du\,dv| = f(x, y)\,|dx\,dy| \tag{35}$$

or

$$g(u, v) = f(x, y)\left|\frac{\partial(x, y)}{\partial(u, v)}\right| = f[\psi_1(u, v), \psi_2(u, v)]\,|J| \tag{36}$$

In (36) the *Jacobian determinant*, or briefly *Jacobian*, is given by

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{37}$$

PROBABILITY DISTRIBUTIONS OF FUNCTIONS OF RANDOM VARIABLES

Theorems 2-2 and 2-4 specifically involve joint probability functions of two random variables. In practice one often needs to find the probability distribution of some specified function of several random variables. Either of the following theorems is often useful for this purpose.

Theorem 2-5: Let X and Y be continuous random variables and let $U = \phi_1(X, Y)$, $V = X$ (the second choice is arbitrary). Then the density function for U is the marginal density obtained from the joint density of U and V as found in Theorem 2-4. A similar result holds for probability functions of discrete variables.

Theorem 2-6: Let $f(x, y)$ be the joint density function of X and Y. Then the density function $g(u)$ of the random variable $U = \phi_1(X, Y)$ is found by differentiating with respect to u the distribution function given by

$$G(u) = P[\phi_1(X, Y) \leqq u] = \iint_{\mathcal{R}} f(x, y)\, dx\, dy \tag{38}$$

where $\mathcal{R}$ is the region for which $\phi_1(x, y) \leqq u$.

CONVOLUTIONS

As a particular consequence of the above theorems we can show (see Problem 2.23) that the density function of the sum of two continuous random variables X and Y, i.e. of $U = X + Y$, having joint density function $f(x, y)$ is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\, dx \tag{39}$$

In the special case where X and Y are independent, $f(x, y) = f_1(x) f_2(y)$ and (39) reduces to

$$g(u) = \int_{-\infty}^{\infty} f_1(x) f_2(u - x)\, dx \tag{40}$$

which is called the *convolution* of f_1 and f_2, abbreviated $f_1 * f_2$.

The following are some important properties of the convolution:

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 * (f_2 + f_3) = f_1 * f_2 + f_1 * f_3$

These results show that f_1, f_2, f_3 obey the *commutative*, *associative* and *distributive laws* of algebra with respect to the operation of convolution.

CONDITIONAL DISTRIBUTIONS

We already know that if $P(A) > 0$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{41}$$

If X and Y are discrete random variables and we have the events $(A: X = x)$, $(B: Y = y)$, then (41) becomes

$$P(Y = y \mid X = x) = \frac{f(x, y)}{f_1(x)} \tag{42}$$

where $f(x, y) = P(X = x, Y = y)$ is the joint probability function and $f_1(x)$ is the marginal probability function for X. We define

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{43}$$

and call it the *conditional probability function of Y given X*. Similarly, the conditional probability function of X given Y is

$$f(x \mid y) \equiv \frac{f(x, y)}{f_2(y)} \tag{44}$$

We shall sometimes denote $f(x \mid y)$ and $f(y \mid x)$ by $f_1(x \mid y)$ and $f_2(y \mid x)$ respectively.

These ideas are easily extended to the case where X, Y are continuous random variables. For example, the *conditional density function of Y given X* is

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{45}$$

where $f(x, y)$ is the joint density function of X and Y, and $f_1(x)$ is the marginal density function of X. Using (45) we can for example find that the probability of Y being between c and d given that $x < X < x + dx$ is

$$P(c < Y < d \mid x < X < x + dx) = \int_c^d f(y \mid x)\, dy \tag{46}$$

Generalizations of these results are also available.

APPLICATIONS TO GEOMETRIC PROBABILITY

Various problems in probability arise from geometric considerations or have geometric interpretations. For example suppose that we have a target in the form of a plane region of area K and a portion of it with area K_1. Then it is reasonable to suppose that the probability of hitting the region of area K_1 is proportional to K_1. We thus define

$$P(\text{hitting region of area } K_1) = \frac{K_1}{K} \tag{47}$$

where it is assumed that the probability of hitting the target is 1. Other assumptions can of course be made. For example there could be less probability of hitting outer areas, etc. The type of assumption used defines the probability distribution function.

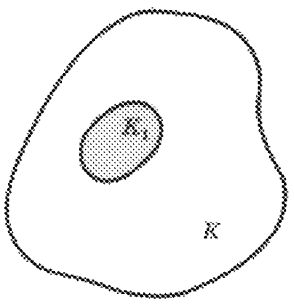


Fig. 2-7

Solved Problems

DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

2.1. Suppose that a pair of fair dice are tossed and let the random variable X denote the sum of the points. (*a*) Obtain the probability distribution for X. (*b*) Construct a graph for this probability distribution.

(*a*) The sample points for tosses of a pair of dice are given in Fig. 1-17, page 18. The random variable X is the sum of the coordinates for each point. Thus for (3, 2) we have $X = 5$. Using the fact that all 36 sample points are equally probable, so that each sample point has probability 1/36, we obtain Table 2-4. For example, corresponding to $X = 5$ we have the four sample points (1, 4), (2, 3), (3, 2), (4, 1), so that the associated probability is 4/36.

Table 2-4

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

(*b*) We can use a bar graph or histogram as given in Fig. 2-8 or Fig. 2-9 according to whether we wish to regard X as a discrete or continuous variable. Note that in Fig. 2-8 the sum of all *ordinates* is 1 while in Fig. 2-9 the sum of all rectangular *areas* is 1.

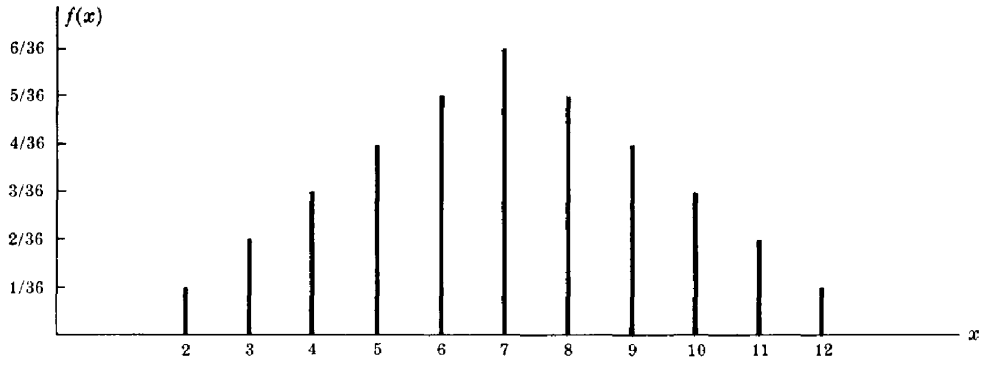


Fig. 2-8

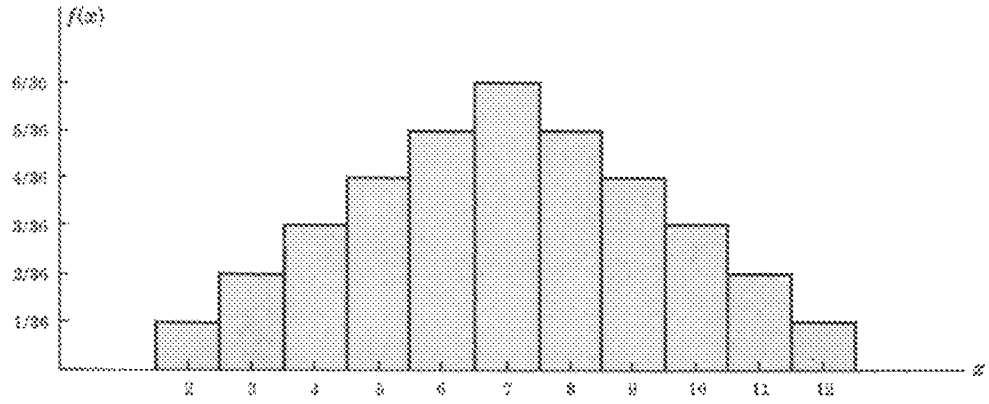


Fig. 2-9

2.2. (*a*) Find the probability distribution of boys and girls in families with 3 children, assuming equal probabilities for boys and girls. (*b*) Graph the distribution in (*a*).

(*a*) Problem 1.46 treated the case of n mutually independent trials, where each trial had just two possible outcomes, A and A', with respective probabilities p and $q = 1 - p$. It was found that the probability of getting exactly x A's in the n trials is ${}_nC_x p^x q^{n-x}$. This result applies to the present problem, under the assumption that successive births (the "trials") are independent so far as the sex of the child is concerned. Thus, with A being the event "a boy," $n = 3$, and $p = q = \frac{1}{2}$, we have

$$P(\text{exactly } x \text{ boys}) = P(X = x) = {}_3C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} = {}_3C_x\left(\frac{1}{2}\right)^3$$

where the random variable X represents the number of boys in the family. (Note that X is defined on the sample space of three trials.) The probability function for X,

$$f(x) = {}_3C_x\left(\frac{1}{2}\right)^3$$

is displayed in Table 2-5.

Table 2-5

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

(*b*) The graph can be represented either as in Fig. 2-10 or Fig. 2-11, depending on whether we wish to consider the variable X as discrete or continuous. Note that the zero of the x-axis has been shifted.

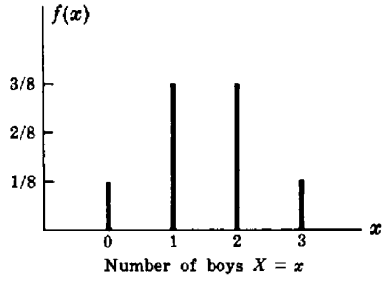


Fig. 2-10

DISCRETE DISTRIBUTION FUNCTIONS

2.3. (*a*) Find the distribution function $F(x)$ for the random variable X of Problem 2.1 and (*b*) graph this distribution function.

(*a*) We have $F(x) = P(X \leqq x) = \sum_{u \leqq x} f(u)$. Then from the results of Problem 2.1 we find

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \leqq x < 3 \\ 3/36 & 3 \leqq x < 4 \\ 6/36 & 4 \leqq x < 5 \\ \vdots & \vdots \\ 35/36 & 11 \leqq x < 12 \\ 1 & 12 \leqq x < \infty \end{cases}$$

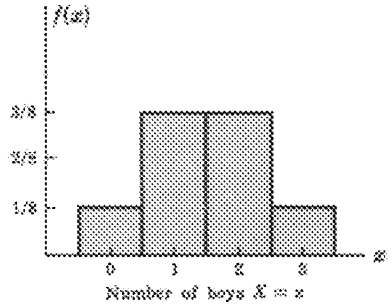


Fig. 2-11

(*b*) See Fig. 2-12.

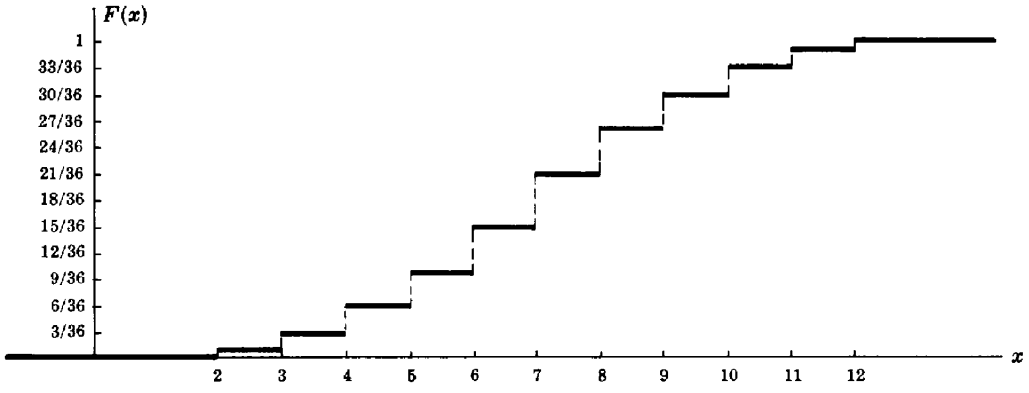


Fig. 2-12

2.4. (*a*) Find the distribution function $F(x)$ for the random variable X of Problem 2.2 and (*b*) graph this distribution function.

(*a*) Using Table 2-5 from Problem 2.2 we obtain

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/8 & 0 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & 3 \leq x < \infty \end{cases}$$

(*b*) The graph of the distribution function of (*a*) is shown in Fig. 2-13.

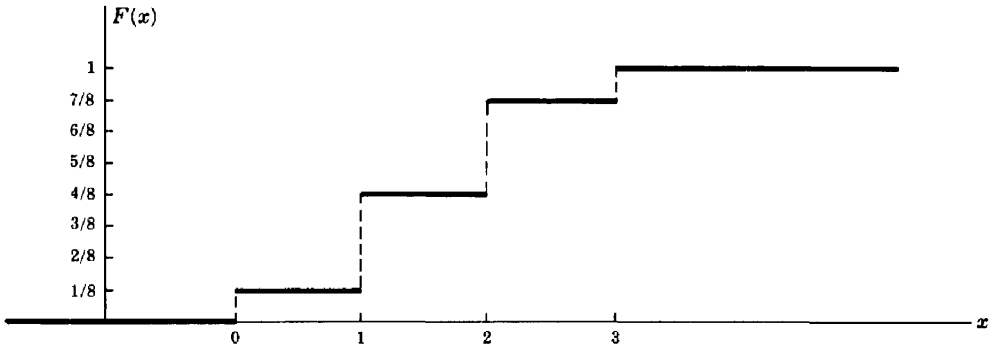


Fig. 2-13

CONTINUOUS RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

2.5. A random variable X has the density function $f(x) = c/(x^2+1)$, where $-\infty < x < \infty$. (*a*) Find the value of the constant c. (*b*) Find the probability that X^2 lies between 1/3 and 1.

(*a*) We must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$, i.e.

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2+1} = c\tan^{-1}x\Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

so that $c = 1/\pi$.

(*b*) If $\frac{1}{3} \leq X^2 \leq 1$, then either $\frac{\sqrt{3}}{3} \leq X \leq 1$ or $-1 \leq X \leq -\frac{\sqrt{3}}{3}$. Thus the required probability is

$$\begin{aligned} \frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2+1} + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2+1} &= \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2+1} \\ &= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right] \\ &= \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6} \end{aligned}$$

2.6. Find the distribution function corresponding to the density function of Problem 2.5.

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(u)\,du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{du}{u^2+1} = \frac{1}{\pi}\left[\tan^{-1}u\Big|_{-\infty}^{x}\right] \\ &= \frac{1}{\pi}[\tan^{-1}x - \tan^{-1}(-\infty)] = \frac{1}{\pi}\left[\tan^{-1}x + \frac{\pi}{2}\right] \\ &= \frac{1}{2} + \frac{1}{\pi}\tan^{-1}x \end{aligned}$$

2.7. The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geqq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leqq 4$.

(a) $$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

(b) $$P(X > 2) = \int_2^\infty 2e^{-2u}\,du = -e^{-2u}\Big|_2^\infty = e^{-4}$$

Another method.

By definition, $P(X \leqq 2) = F(2) = 1 - e^{-4}$. Hence,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

(c) $$P(-3 < X \leqq 4) = \int_{-3}^4 f(u)\,du = \int_{-3}^0 0\,du + \int_0^4 2e^{-2u}\,du$$

$$= -e^{-2u}\Big|_0^4 = 1 - e^{-8}$$

Another method.

$$P(-3 < X \leqq 4) = P(X \leqq 4) - P(X \leqq -3)$$
$$= F(4) - F(-3)$$
$$= (1 - e^{-8}) - (0) = 1 - e^{-8}$$

JOINT DISTRIBUTIONS AND INDEPENDENT VARIABLES

2.8. The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \leqq x \leqq 2$, $0 \leqq y \leqq 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c.

(b) Find $P(X = 2, Y = 1)$.

(c) Find $P(X \geqq 1, Y \leqq 2)$.

(a) The sample points (x, y) for which probabilities are different from zero are indicated in Fig. 2-14. The probabilities associated with these points, given by $c(2x + y)$, are shown in Table 2-6. Since the grand total, $42c$, must equal 1, we have $c = 1/42$.

Table 2-6

X \ Y	0	1	2	3	Totals ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totals →	$6c$	$9c$	$12c$	$15c$	$42c$

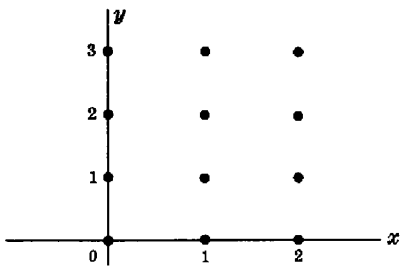


Fig. 2-14

(*b*) From Table 2-6 we see that

$$P(X=2, Y=1) = 5c = \frac{5}{42}$$

(*c*) From Table 2-6 we see that

$$P(X \geqq 1, Y \leqq 2) = \sum_{x \geqq 1} \sum_{y \leqq 2} f(x,y)$$
$$= (2c+3c+4c) + (4c+5c+6c)$$
$$= 24c = \frac{24}{42} = \frac{4}{7}$$

as indicated by the entries shown shaded in the table.

2.9. Find the marginal probability functions (*a*) of X and (*b*) of Y for the random variables of Problem 2.8.

(*a*) The marginal probability function for X is given by $P(X=x) = f_1(x)$ and can be obtained from the margin totals in the right-hand column of Table 2-6. From these we see that

$$P(X=x) = f_1(x) = \begin{cases} 6c = 1/7 & x = 0 \\ 14c = 1/3 & x = 1 \\ 22c = 11/21 & x = 2 \end{cases}$$

Check: $\frac{1}{7} + \frac{1}{3} + \frac{11}{21} = 1$

(*b*) The marginal probability function for Y is given by $P(Y=y) = f_2(y)$ and can be obtained from the margin totals in the last row of Table 2-6. From these we see that

$$P(Y=y) = f_2(y) = \begin{cases} 6c = 1/7 & y = 0 \\ 9c = 3/14 & y = 1 \\ 12c = 2/7 & y = 2 \\ 15c = 5/14 & y = 3 \end{cases}$$

Check: $\frac{1}{7} + \frac{3}{14} + \frac{2}{7} + \frac{5}{14} = 1$

2.10. Show that the random variables X and Y of Problem 2.8 are dependent.

If the random variables X and Y are independent then we must have, for all x and y,

$$P(X=x, Y=y) = P(X=x)\,P(Y=y)$$

But, as seen from Problems 2.8(*b*) and 2.9,

$$P(X=2, Y=1) = \frac{5}{42} \qquad P(X=2) = \frac{11}{21} \qquad P(Y=1) = \frac{3}{14}$$

so that

$$P(X=2, Y=1) \neq P(X=2)\,P(Y=1)$$

The result also follows from the fact that the joint probability function, $(2x+y)/42$, cannot be expressed as a function of x alone times a function of y alone.

2.11. The joint density function of two continuous random variables X and Y is

$$f(x,y) = \begin{cases} cxy & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Find the value of the constant c.

(*b*) Find $P(1 < X < 2,\ 2 < Y < 3)$.

(*c*) Find $P(X \geqq 3, Y \leqq 2)$.

(*a*) We must have the total probability equal to 1, i.e.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

Ziyad Asaad

Using the definition of $f(x, y)$, the integral has the value

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\,dx\,dy = c\int_{x=0}^{4}\left[\int_{y=1}^{5} xy\,dy\right]dx$$

$$= c\int_{x=0}^{4}\frac{xy^2}{2}\Big|_{y=1}^{5}dx = c\int_{x=0}^{4}\left(\frac{25x}{2}-\frac{x}{2}\right)dx$$

$$= c\int_{x=0}^{4} 12x\,dx = c(6x^2)\Big|_{x=0}^{4} = 96c$$

Then $96c = 1$ and $c = 1/96$.

(*b*) Using the value of c found in (*a*) we have

$$P(1 < X < 2,\ 2 < Y < 3) = \int_{x=1}^{2}\int_{y=2}^{3}\frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,dy\right]dx = \frac{1}{96}\int_{x=1}^{2}\frac{xy^2}{2}\Big|_{y=2}^{3}dx$$

$$= \frac{1}{96}\int_{x=1}^{2}\frac{5x}{2}\,dx = \frac{5}{192}\left(\frac{x^2}{2}\right)\Big|_{1}^{2} = \frac{5}{128}$$

(*c*) $$P(X \geq 3,\ Y \leq 2) = \int_{x=3}^{4}\int_{y=1}^{2}\frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=3}^{4}\left[\int_{y=1}^{2} xy\,dy\right]dx = \frac{1}{96}\int_{x=3}^{4}\frac{xy^2}{2}\Big|_{y=1}^{2}dx$$

$$= \frac{1}{96}\int_{x=3}^{4}\frac{3x}{2}\,dx = \frac{7}{128}$$

2.12. Find the marginal distribution functions (*a*) of X and (*b*) of Y for Problem 2.11.

(*a*) The marginal distribution function for X if $0 \leq x < 4$ is

$$F_1(x) = P(X \leq x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{x}\int_{v=1}^{5}\frac{uv}{96}\,du\,dv$$

$$= \frac{1}{96}\int_{u=0}^{x}\left[\int_{v=1}^{5} uv\,dv\right]du = \frac{x^2}{16}$$

For $x \geq 4$, $F_1(x) = 1$; for $x < 0$, $F_1(x) = 0$. Thus

$$F_1(x) = \begin{cases} 0 & x < 0 \\ x^2/16 & 0 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

As $F_1(x)$ is continuous at $x = 0$ and $x = 4$, we could replace $<$ by $\leq$ in the above expression.

(*b*) The marginal distribution function for Y if $1 \leq y < 5$ is

$$F_2(y) = P(Y \leq y) = \int_{u=-\infty}^{\infty}\int_{v=1}^{y} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{4}\int_{v=1}^{y}\frac{uv}{96}\,du\,dv = \frac{y^2-1}{24}$$

For $y \geqq 5$, $F_2(y) = 1$. For $y < 1$, $F_2(y) = 0$. Thus

$$F_2(y) = \begin{cases} 0 & y < 1 \\ (y^2 - 1)/24 & 1 \leqq y < 5 \\ 1 & y \geqq 5 \end{cases}$$

As $F_2(y)$ is continuous at $y = 1$ and $y = 5$, we could replace $<$ by $\leqq$ in the above expression.

2.13. Find the joint distribution function for the random variables X, Y of Problem 2.11.

From Problem 2.11 it is seen that the joint density function for X and Y can be written as the product of a function of x alone and a function of y alone. In fact, $f(x, y) = f_1(x) f_2(y)$, where

$$f_1(x) = \begin{cases} c_1 x & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases} \qquad f_2(y) = \begin{cases} c_2 y & 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

and $c_1 c_2 = c = 1/96$. It follows that X and Y are independent, so that their joint distribution function is given by $F(x, y) = F_1(x) F_2(y)$. The marginal distributions $F_1(x)$ and $F_2(y)$ were determined in Problem 2.12, and Fig. 2-15 shows the resulting piecewise definition of $F(x, y)$.

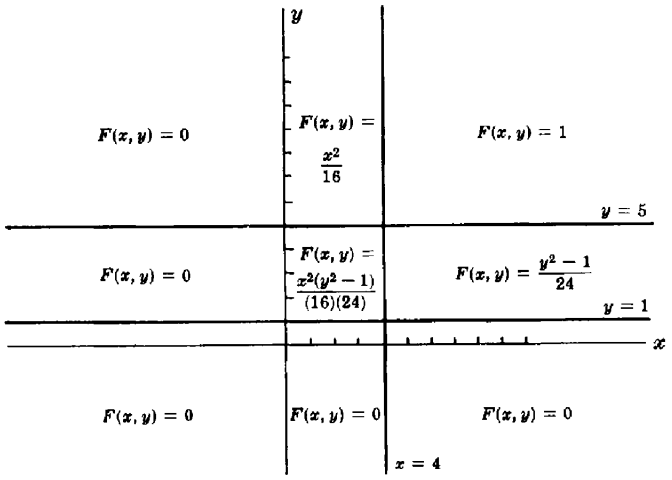


Fig. 2-15

2.14. In Problem 2.11 find $P(X + Y < 3)$.

In Fig. 2-16 we have indicated the square region $0 < x < 4$, $1 < y < 5$ within which the joint density function of X and Y is different from zero. The required probability is given by

$$P(X + Y < 3) = \iint_{\mathcal{R}} f(x, y)\, dx\, dy$$

where $\mathcal{R}$ is the part of the square over which $x + y < 3$, shown shaded in Fig. 2-16. Since $f(x, y) = xy/96$ over $\mathcal{R}$, this probability is given by

$$\int_{x=0}^{2} \int_{y=1}^{3-x} \frac{xy}{96}\, dx\, dy$$

$$= \frac{1}{96} \int_{x=0}^{2} \left[\int_{y=1}^{3-x} xy\, dy \right] dx$$

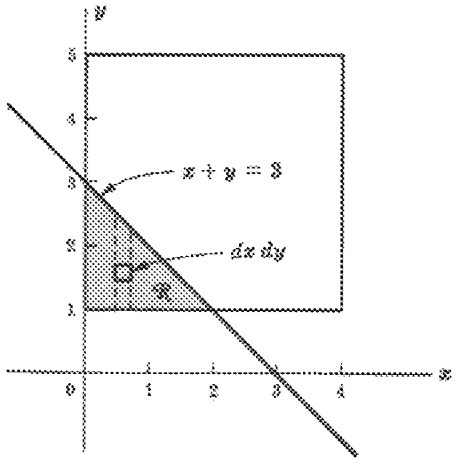


Fig. 2-16

$$= \frac{1}{96}\int_{x=0}^{2} \frac{xy^2}{2}\Big|_{y=1}^{3-x} dx = \frac{1}{192}\int_{x=0}^{2} [x(3-x)^2 - x]\,dx = \frac{1}{48}$$

CHANGE OF VARIABLES

2.15. Prove Theorem 2-1, page 46.

The probability function for U is given by

$$g(u) = P(U = u) = P[\phi(X) = u] = P[X = \psi(u)] = f[\psi(u)]$$

In a similar manner Theorem 2-2, page 46, can be proved. See Problem 2.66.

2.16. Prove Theorem 2-3, page 46.

Consider first the case where $u = \phi(x)$ or $x = \psi(u)$ is an increasing function, i.e. u increases as x increases (Fig. 2-17). Then as is clear from the figure we have

(1) $$P(u_1 < U < u_2) = P(x_1 < X < x_2)$$

or

(2) $$\int_{u_1}^{u_2} g(u)\,du = \int_{x_1}^{x_2} f(x)\,dx$$

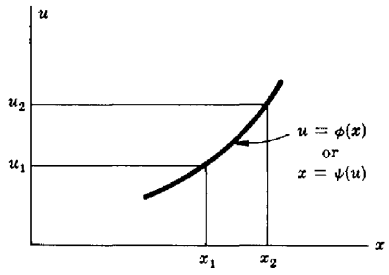


Fig. 2-17

Letting $x = \psi(u)$ in the integral on the right, (2) can be written

$$\int_{u_1}^{u_2} g(u)\,du = \int_{u_1}^{u_2} f[\psi(u)]\psi'(u)\,du$$

This can hold for all u_1 and u_2 only if the integrands are identical, i.e.

$$g(u) = f[\psi(u)]\psi'(u)$$

This is a special case of (34), page 46, where $\psi'(u) > 0$ (i.e. the slope is positive). For the case where $\psi'(u) \leqq 0$, i.e. u is a decreasing function of x, we can also show that (34) holds (see Problem 2.67). The theorem can also be proved if $\psi'(u) \geqq 0$ or $\psi'(u) < 0$.

2.17. Prove Theorem 2-4, page 46.

We suppose first that as x and y increase, u and v also increase. As in Problem 2.16 we can then show that

$$P(u_1 < U < u_2,\ v_1 < V < v_2) = P(x_1 < X < x_2,\ y_1 < Y < y_2)$$

or

$$\int_{u_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x, y)\,dx\,dy$$

Letting $x = \psi_1(u, v)$, $y = \psi_2(u, v)$ in the integral on the right we have by a theorem of advanced calculus

$$\int_{u_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{u_1}^{u_2}\int_{v_1}^{v_2} f[\psi_1(u, v), \psi_2(u, v)]J\,du\,dv$$

where

$$J = \frac{\partial(x, y)}{\partial(u, v)}$$

is the *Jacobian*. Thus

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)]J$$

which is (36), page 46, in the case where $J > 0$. Similarly we can prove (36) for the case where $J < 0$.

2.18. The probability function of a random variable X is

$$f(x) = \begin{cases} 2^{-x} & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Find the probability function for the random variable $U = X^4 + 1$.

Since $U = X^4 + 1$ the relationship between the values u and x of the random variables U and X is given by $u = x^4 + 1$ or $x = \sqrt[4]{u-1}$, where $u = 2, 17, 82, \ldots$ and the real positive root is taken. Then the required probability function for U is given by

$$g(u) = \begin{cases} 2^{-\sqrt[4]{u-1}} & u = 2, 17, 82, \ldots \\ 0 & \text{otherwise} \end{cases}$$

using Theorem 2-1, page 46, or Problem 2.15.

2.19. The probability function of a random variable X is given by

$$f(x) = \begin{cases} x^2/81 & -3 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density for the random variable $U = \frac{1}{3}(12 - X)$.

We have $u = \frac{1}{3}(12 - x)$ or $x = 12 - 3u$. Thus to each value of x there is one and only one value of u and conversely. The values of u corresponding to $x = -3$ and $x = 6$ are $u = 5$ and $u = 2$ respectively. Since $\psi'(u) = dx/du = -3$ it follows by Theorem 2-3, page 46, or Problem 2.16 that the density function for U is

$$g(u) = \begin{cases} (12 - 3u)^2/27 & 2 < u < 5 \\ 0 & \text{otherwise} \end{cases}$$

Check:
$$\int_2^5 \frac{(12-3u)^2}{27}\,du = -\frac{(12-3u)^3}{243}\Big|_2^5 = 1$$

2.20. Find the probability density of the random variable $U = X^2$ where X is the random variable of Problem 2.19.

We have $u = x^2$ or $x = \pm\sqrt{u}$. Thus to each value of x there corresponds one and only one value of u but to each value of $u \neq 0$ there correspond *two* values of x. The values of x for which $-3 < x < 6$ correspond to values of u for which $0 \leqq u < 36$ as shown in Fig. 2-18.

As seen in this figure the interval $-3 < x \leqq 3$ corresponds to $0 \leqq u \leqq 9$ while $3 < x < 6$ corresponds to $9 < u < 36$. In this case we cannot use Theorem 2-3 directly but can proceed as follows. The distribution function for U is

$$G(u) = P(U \leqq u)$$

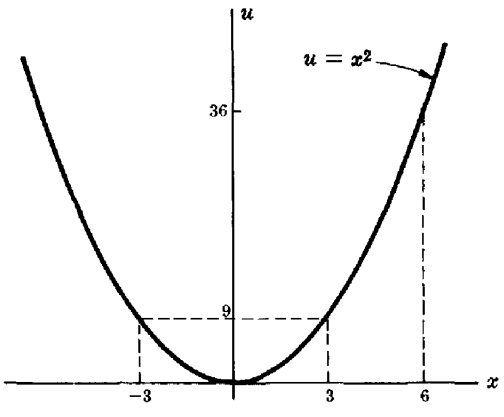


Fig. 2-18

Now if $0 \leqq u \leqq 9$ we have

$$G(u) = P(U \leqq u) = P(X^2 \leqq u) = P(-\sqrt{u} \leqq X \leqq \sqrt{u}) = \int_{-\sqrt{u}}^{\sqrt{u}} f(x)\,dx$$

But if $9 < u < 36$ we have

$$G(u) = P(U \leqq u) = P(-3 < X < \sqrt{u}) = \int_{-3}^{\sqrt{u}} f(x)\,dx$$

Since the density function $g(u)$ is the derivative of $G(u)$ we have, using Leibniz's rule (*12*),

$$g(u) = \begin{cases} \dfrac{f(\sqrt{u}) + f(-\sqrt{u})}{2\sqrt{u}} & 0 \leqq u \leqq 9 \\ \dfrac{f(\sqrt{u})}{2\sqrt{u}} & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Using the given definition of $f(x)$ this becomes

$$g(u) = \begin{cases} \sqrt{u}/81 & 0 \leqq u \leqq 9 \\ \sqrt{u}/162 & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Check: $$\int_0^9 \frac{\sqrt{u}}{81}du + \int_9^{36} \frac{\sqrt{u}}{162}du = \frac{2u^{3/2}}{243}\Big|_0^9 + \frac{u^{3/2}}{243}\Big|_9^{36} = 1$$

2.21. If the random variables X and Y have joint density function

$$f(x,y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(see Problem 2.11) find the density function of $U = X + 2Y$.

Method 1.

Let $u = x + 2y$, $v = x$, the second relation being chosen arbitrarily. Then simultaneous solution yields $x = v$, $y = \frac{1}{2}(u - v)$. Thus the region $0 < x < 4$, $1 < y < 5$ corresponds to the region $0 < v < 4$, $2 < u - v < 10$ shown shaded in Fig. 2-19.

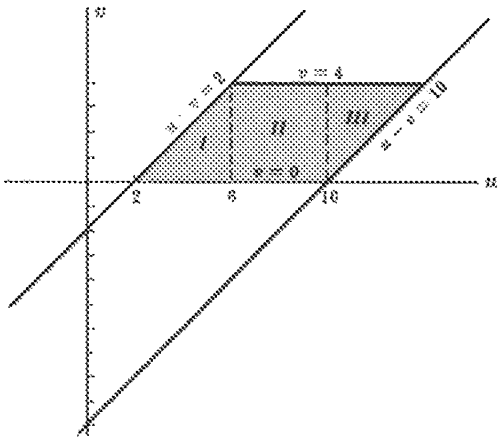


Fig. 2-19

The Jacobian is given by

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Then by Theorem 2-4 the joint density function of U and V is

$$g(u,v) = \begin{cases} v(u-v)/384 & 2 < u - v < 10,\ 0 < v < 4 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function of U is given by

$$g_1(u) = \begin{cases} \displaystyle\int_{v=0}^{u-2} \frac{v(u-v)}{384}dv & 2 < u < 6 \\ \displaystyle\int_{v=0}^{4} \frac{v(u-v)}{384}dv & 6 < u < 10 \\ \displaystyle\int_{v=u-10}^{4} \frac{v(u-v)}{384}dv & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

as seen by referring to the shaded regions *I, II, III* of Fig. 2-19. Carrying out the integrations we find

$$g_1(u) = \begin{cases} (u-2)^2(u+4)/2304 & 2 < u < 6 \\ (3u-8)/144 & 6 < u < 10 \\ (348u - u^3 - 2128)/2304 & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

A check can be achieved by showing that the integral of $g_1(u)$ is equal to 1.

Method 2.

The distribution function of the random variable $X + 2Y$ is given by

$$P(X + 2Y \leqq u) = \iint_{x+2y \leqq u} f(x, y)\, dx\, dy = \iint_{\substack{x+2y \leqq u \\ 0 < x < 4 \\ 1 < y < 5}} \frac{xy}{96}\, dx\, dy$$

For $2 < u < 6$ we see by referring to Fig. 2-20 that the last integral equals

$$\int_{x=0}^{u-2} \int_{y=1}^{(u-x)/2} \frac{xy}{96}\, dx\, dy = \int_{x=0}^{u-2} \left[\frac{x(u-x)^2}{768} - \frac{x}{192} \right] dx$$

The derivative of this with respect to u is found to be $(u-2)^2(u+4)/2304$. In a similar manner we can obtain the result of Method 1 for $6 < u < 10$, etc.

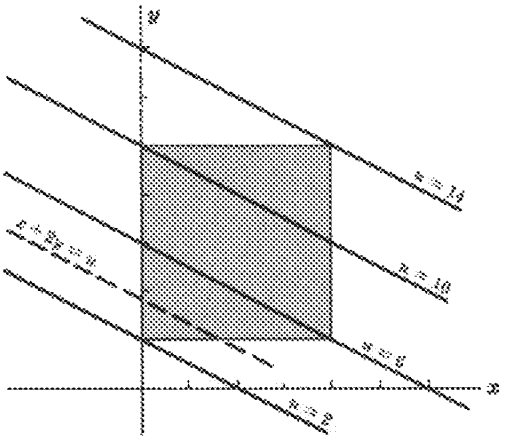


Fig. 2-20

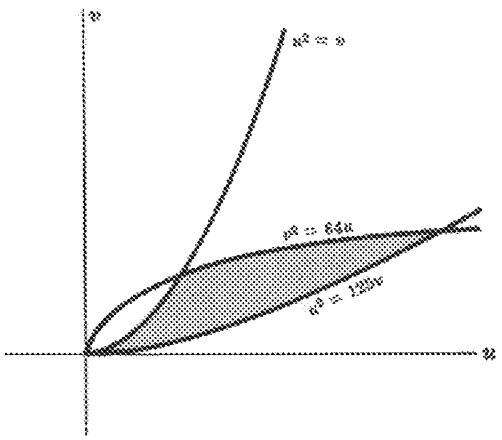


Fig. 2-21

2.22. If the random variables X and Y have joint density function

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(see Problem 2.11) find the joint density function of $U = XY^2$, $V = X^2Y$.

Consider $u = xy^2$, $v = x^2y$. Dividing these equations we obtain $y/x = u/v$ so that $y = ux/v$. This leads to the simultaneous solution $x = v^{2/3}u^{-1/3}$, $y = u^{2/3}v^{-1/3}$. The image of $0 < x < 4$, $1 < y < 5$ in the uv-plane is given by

$$0 < v^{2/3}u^{-1/3} < 4 \qquad 1 < u^{2/3}v^{-1/3} < 5$$

which are equivalent to

$$v^2 < 64u \qquad v < u^2 < 125v$$

This region is shown shaded in Fig. 2-21.

The Jacobian is given by

$$J = \begin{vmatrix} -\frac{1}{3}v^{2/3}u^{-4/3} & \frac{2}{3}v^{-1/3}u^{-1/3} \\ \frac{2}{3}u^{-1/3}v^{-1/3} & -\frac{1}{3}u^{2/3}v^{-4/3} \end{vmatrix} = -\frac{1}{3}u^{-2/3}v^{-2/3}$$

Thus the joint density function of U and V is, by Theorem 2-4,

$$g(u,v) = \begin{cases} \frac{(v^{2/3}u^{-1/3})(u^{2/3}v^{-1/3})}{96}(\frac{1}{3}u^{-2/3}v^{-2/3}) & v^2 < 64u,\ v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

or

$$g(u,v) = \begin{cases} u^{-1/3}v^{-1/3}/288 & v^2 < 64u,\ v < u^2 < 125v \\ 0 & \text{otherwise} \end{cases}$$

CONVOLUTIONS

2.23. Let X and Y be random variables having joint density function $f(x,y)$. Prove that the density function of $U = X + Y$ is

$$g(u) = \int_{-\infty}^{\infty} f(v, u-v)\,dv$$

Method 1.

Let $U = X + Y$, $V = X$, where we have arbitrarily added the second equation. Corresponding to these we have $u = x + y$, $v = x$ or $x = v$, $y = u - v$. The Jacobian of the transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

Thus by Theorem 2-4, page 46, the joint density function of U and V is

$$g(u,v) = f(v, u-v)$$

It follows from (*26*), page 45, that the marginal density function of U is

$$g(u) = \int_{-\infty}^{\infty} f(v, u-v)\,dv$$

Method 2.

The distribution function of $U = X + Y$ is equal to the double integral of $f(x,y)$ taken over the region defined by $x + y \leqq u$, i.e.

$$G(u) = \iint_{x+y \leqq u} f(x,y)\,dx\,dy$$

Since the region is below the line $x + y = u$, as indicated by the shading in Fig. 2-22, we see that

$$G(u) = \int_{x=-\infty}^{\infty}\left[\int_{y=-\infty}^{u-x} f(x,y)\,dy\right]dx$$

The density function of U is the derivative of $G(u)$ with respect to u and is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u-x)\,dx$$

using Leibniz's rule (*12*) first on the x-integral and then on the y-integral.

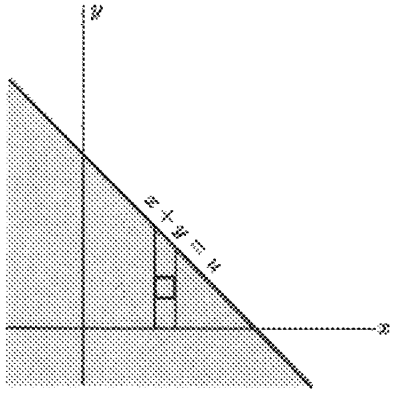


Fig. 2-22

2.24. Work Problem 2.23 if X and Y are independent random variables having density functions $f_1(x)$, $f_2(y)$ respectively.

In this case the joint density function is $f(x, y) = f_1(x) f_2(y)$, so that by Problem 2.23 the density function of $U = X + Y$ is

$$g(u) = \int_{-\infty}^{\infty} f_1(v) f_2(u - v)\, dv = f_1 * f_2$$

which is the *convolution* of f_1 and f_2.

2.25. If X and Y are independent random variables having density functions

$$f_1(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad f_2(y) = \begin{cases} 3e^{-3y} & y \geq 0 \\ 0 & y < 0 \end{cases}$$

find the density function of their sum, $U = X + Y$.

By Problem 2.24 the required density function is the convolution of f_1 and f_2 and is given by

$$g(u) = f_1 * f_2 = \int_{-\infty}^{\infty} f_1(v) f_2(u - v)\, dv$$

In the integrand f_1 vanishes when $v < 0$ and f_2 vanishes when $v > u$. Hence

$$g(u) = \int_0^u (2e^{-2v})(3e^{-3(u-v)})\, dv$$

$$= 6e^{-3u} \int_0^u e^v\, dv = 6e^{-3u}(e^u - 1) = 6(e^{-2u} - e^{-3u})$$

if $u \geq 0$ and $g(u) = 0$ if $u < 0$.

Check: $$\int_{-\infty}^{\infty} g(u)\, du = 6 \int_0^{\infty} (e^{-2u} - e^{-3u})\, du = 6\left(\frac{1}{2} - \frac{1}{3}\right) = 1$$

2.26. Prove that $f_1 * f_2 = f_2 * f_1$ (Property 1, page 47).

We have

$$f_1 * f_2 = \int_{v=-\infty}^{\infty} f_1(v) f_2(u - v)\, dv$$

Letting $w = u - v$ so that $v = u - w$, $dv = -dw$, we obtain

$$f_1 * f_2 = \int_{w=\infty}^{-\infty} f_1(u - w) f_2(w)(-dw) = \int_{w=-\infty}^{\infty} f_2(w) f_1(u - w)\, dw = f_2 * f_1$$

CONDITIONAL DISTRIBUTIONS

2.27. Find (*a*) $f(y \mid 2)$, (*b*) $P(Y = 1 \mid X = 2)$ for the distribution of Problem 2.8.

(*a*) Using the results in Problems 2.8 and 2.9 we have

$$f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{(2x + y)/42}{f_1(x)}$$

so that with $x = 2$

$$f(y \mid 2) = \frac{(4 + y)/42}{11/21} = \frac{4 + y}{22}$$

(*b*) $$P(Y = 1 \mid X = 2) = f(1 \mid 2) = \frac{5}{22}$$

2.28. If X and Y have the joint density function

$$f(x,y) = \begin{cases} \frac{3}{4} + xy & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

find (*a*) $f(y \mid x)$, (*b*) $P(Y > \frac{1}{2} \mid \frac{1}{2} < X < \frac{1}{2} + dx)$.

(*a*) For $0 < x < 1$,

$$f_1(x) = \int_0^1 \left(\frac{3}{4} + xy\right) dy = \frac{3}{4} + \frac{x}{2}$$

and

$$f(y \mid x) = \frac{f(x,y)}{f_1(x)} = \begin{cases} \dfrac{3 + 4xy}{3 + 2x} & 0 < y < 1 \\ 0 & \text{other } y \end{cases}$$

For other values of x, $f(y \mid x)$ is not defined.

(*b*) $$P(Y > \tfrac{1}{2} \mid \tfrac{1}{2} < X < \tfrac{1}{2} + dx) = \int_{1/2}^{\infty} f(y \mid \tfrac{1}{2})\, dy = \int_{1/2}^{1} \frac{3 + 2y}{4}\, dy = \frac{9}{16}$$

2.29. The joint density function of the random variables X and Y is given by

$$f(x,y) = \begin{cases} 8xy & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq x \\ 0 & \text{otherwise} \end{cases}$$

Find (*a*) the marginal density of X, (*b*) the marginal density of Y, (*c*) the conditional density of X, (*d*) the conditional density of Y.

The region over which $f(x,y)$ is different from zero is shown shaded in Fig. 2-23.

(*a*) To obtain the marginal density of X we fix x and integrate with respect to y from 0 to x as indicated by the vertical strip in Fig. 2-23. The result is

$$f_1(x) = \int_{y=0}^{x} 8xy\, dy = 4x^3$$

for $0 < x < 1$. For all other values of x, $f_1(x) = 0$.

(*b*) Similarly, the marginal density of Y is obtained by fixing y and integrating with respect to x from $x = y$ to $x = 1$, as indicated by the horizontal strip in Fig. 2-23. The result is, for $0 < y < 1$,

$$f_2(y) = \int_{x=y}^{1} 8xy\, dx = 4y(1 - y^2)$$

For all other values of y, $f_2(y) = 0$.

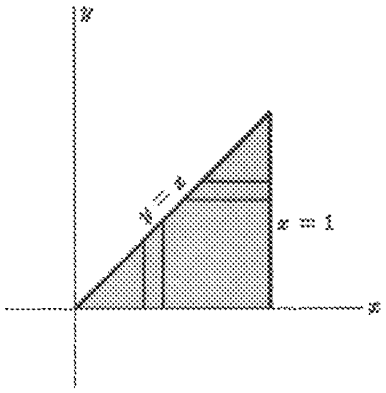


Fig. 2-23

(*c*) The conditional density function of X is, for $0 < y < 1$,

$$f_1(x \mid y) = \frac{f(x,y)}{f_2(y)} = \begin{cases} 2x/(1 - y^2) & y \leqq x \leqq 1 \\ 0 & \text{other } x \end{cases}$$

The conditional density function is not defined when $f_2(y) = 0$.

(*d*) The conditional density function of Y is, for $0 < x < 1$,

$$f_2(y \mid x) = \frac{f(x,y)}{f_1(x)} = \begin{cases} 2y/x^2 & 0 \leqq y \leqq x \\ 0 & \text{other } y \end{cases}$$

The conditional density function is not defined when $f_1(x) = 0$.

Check: $\int_0^1 f_1(x)\,dx = \int_0^1 4x^3\,dx = 1, \qquad \int_0^1 f_2(y)\,dy = \int_0^1 4y(1-y^2)\,dy = 1$

$$\int_y^1 f_1(x \mid y)\,dx = \int_y^1 \frac{2x}{1-y^2}\,dx = 1$$

$$\int_0^x f_2(y \mid x)\,dy = \int_0^x \frac{2y}{x^2}\,dy = 1$$

2.30. Determine whether the random variables of Problem 2.29 are independent.

In the shaded region of Fig. 2-23 $f(x,y) = 8xy$, $f_1(x) = 4x^3$, $f_2(y) = 4y(1-y^2)$. Hence $f(x,y) \neq f_1(x)\,f_2(y)$, and thus X and Y are dependent.

It should be noted that it does not follow from $f(x,y) = 8xy$ that $f(x,y)$ can be expressed as a function of x alone times a function of y alone. This is because the restriction $0 \leqq y \leqq x$ occurs. If this were replaced by some restriction on y not depending on x (as in Problem 2.21), such a conclusion would be valid.

APPLICATIONS TO GEOMETRIC PROBABILITY

2.31. A person playing darts finds that the probability of the dart striking between r and $r + dr$ is

$$P(r \leqq R \leqq r+dr) = c\left[1 - \left(\frac{r}{a}\right)^2\right]dr$$

Here, R is the distance of the hit from the center of the target, c is a constant, and a is the radius of the target (see Fig. 2-24). Find the probability of hitting the bull's-eye, which is assumed to have radius b. Assume that the target is always hit.

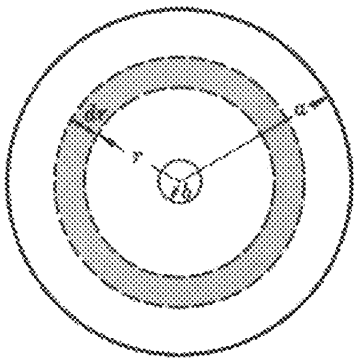


Fig. 2-24

The density function is given by

$$f(r) = c\left[1 - \left(\frac{r}{a}\right)^2\right]$$

Since the target is always hit, we have

$$c\int_0^a \left[1 - \left(\frac{r}{a}\right)^2\right]dr = 1$$

from which $c = 3/2a$. Then the probability of hitting the bull's-eye is

$$\int_0^b f(r)\,dr = \frac{3}{2a}\int_0^b \left[1 - \left(\frac{r}{a}\right)^2\right]dr = \frac{b(3a^2 - b^2)}{2a^3}$$

2.32. Two points are selected at random in the interval $0 \leqq x \leqq 1$. Determine the probability that the sum of their squares is less than 1.

Let X and Y denote the random variables associated with the given points. Since equal intervals are assumed to have equal probabilities, the density functions of X and Y are given respectively by

(1) $$f_1(x) = \begin{cases} 1 & 0 \leqq x \leqq 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_2(y) = \begin{cases} 1 & 0 \leqq y \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

Then since X and Y are independent, the joint density function is given by

$$(2) \qquad f(x,y) = f_1(x)\,f_2(y) = \begin{cases} 1 & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

It follows that the required probability is given by

$$(3) \qquad P(X^2 + Y^2 \leqq 1) = \iint_{\mathcal{R}} dx\,dy$$

where $\mathcal{R}$ is the region defined by $x^2 + y^2 \leqq 1$, $x \geqq 0$, $y \geqq 0$, which is a quarter of a circle of radius 1 (Fig. 2-25). Now since (3) represents the area of $\mathcal{R}$ we see that the required probability is $\pi/4$.

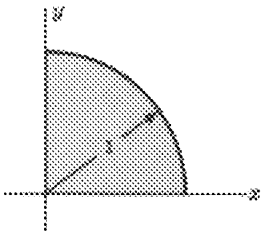


Fig. 2-25

MISCELLANEOUS PROBLEMS

2.33. Suppose that the random variables X and Y have joint density function given by

$$f(x,y) = \begin{cases} c(2x+y) & 2 < x < 6,\ 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

Find (*a*) the constant c, (*b*) the marginal distribution functions for X and Y, (*c*) the marginal density functions for X and Y, (*d*) $P(3 < X < 4, Y > 2)$, (*e*) $P(X > 3)$, (*f*) $P(X + Y > 4)$, (*g*) the joint distribution function, (*h*) whether X and Y are independent.

(*a*) The total probability is given by

$$\int_{x=2}^{6}\int_{y=0}^{5} c(2x+y)\,dx\,dy = \int_{x=2}^{6} c\left(2xy + \frac{y^2}{2}\right)\Big|_0^5 dx$$

$$= \int_{x=2}^{6} c\left(10x + \frac{25}{2}\right)dx = 210c$$

For this to equal 1 we must have $c = 1/210$.

(*b*) The marginal distribution function for X is

$$F_1(x) = P(X \leqq x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u,v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} 0\,du\,dv = 0 & x < 2 \\ \displaystyle\int_{u=2}^{x}\int_{v=0}^{5} \frac{2u+v}{210}\,du\,dv = \frac{2x^2 + 5x - 18}{84} & 2 \leqq x < 6 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u+v}{210}\,du\,dv = 1 & x \geqq 6 \end{cases}$$

The marginal distribution function for Y is

$$F_2(y) = P(Y \leqq y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u,v)\,du\,dv$$

$$= \begin{cases} \displaystyle\int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} 0\,du\,dv = 0 & y < 0 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{y} \frac{2u+v}{210}\,du\,dv = \frac{y^2 + 16y}{105} & 0 \leqq y < 5 \\ \displaystyle\int_{u=2}^{6}\int_{v=0}^{5} \frac{2u+v}{210}\,du\,dv = 1 & y \geqq 5 \end{cases}$$

(*c*) The marginal density function for X is, from part (*b*),

$$f_1(x) = \frac{d}{dx}F_1(x) = \begin{cases} (4x+5)/84 & 2 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function for Y is, from part (*b*),

$$f_2(y) = \frac{d}{dy}F_2(y) = \begin{cases} (2y+16)/105 & 0 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(*d*) $$P(3 < X < 4,\ Y > 2) = \frac{1}{210}\int_{x=3}^{4}\int_{y=2}^{5}(2x+y)\,dx\,dy = \frac{3}{20}$$

(*e*) $$P(X > 3) = \frac{1}{210}\int_{x=3}^{6}\int_{y=0}^{5}(2x+y)\,dx\,dy = \frac{23}{28}$$

(*f*) $$P(X+Y > 4) = \iint_{\mathcal{R}} f(x,y)\,dx\,dy$$

where $\mathcal{R}$ is the shaded region of Fig. 2-26. Although this can be found, it is easier to use the fact that

$$P(X+Y > 4) = 1 - P(X+Y \leqq 4) = 1 - \iint_{\mathcal{R}'} f(x,y)\,dx\,dy$$

where $\mathcal{R}'$ is the cross hatched region of Fig. 2-26. We have

$$P(X+Y \leqq 4) = \frac{1}{210}\int_{x=2}^{4}\int_{y=0}^{4-x}(2x+y)\,dx\,dy = \frac{2}{35}$$

Thus $P(X+Y > 4) = 33/35$.

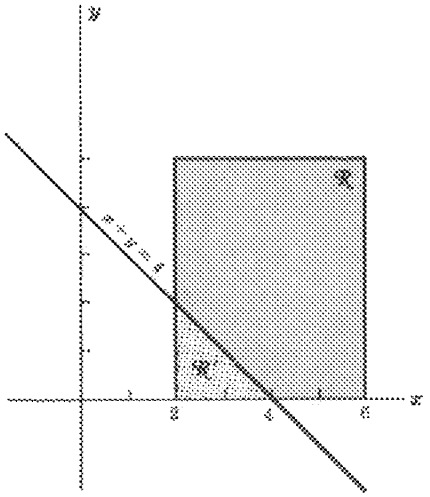


Fig. 2-26

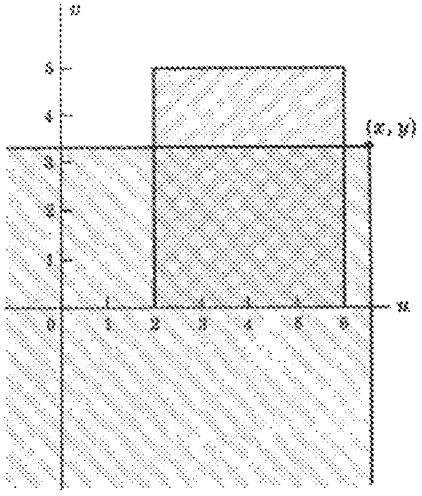


Fig. 2-27

(*g*) The joint distribution function is

$$F(x,y) = P(X \leqq x,\ Y \leqq y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u,v)\,du\,dv$$

In the uv-plane (Fig. 2-27) the region of integration is the intersection of the quarter-plane $u \leqq x$, $v \leqq y$ and the rectangle $2 < u < 6$, $0 < v < 5$ [over which $f(u,v)$ is nonzero]. For (x,y) located as in the figure, we have

$$F(x,y) = \int_{u=2}^{6}\int_{v=0}^{y}\frac{2u+v}{210}\,du\,dv = \frac{16y+y^2}{105}$$

When (x, y) lies inside the rectangle, we obtain another expression, etc. The complete results are shown in Fig. 2-28.

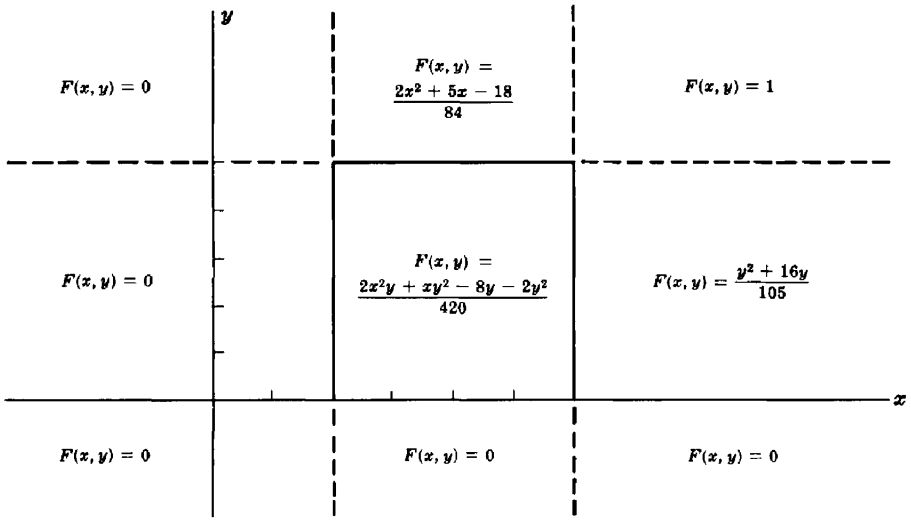


Fig. 2-28

(*h*) The random variables are dependent since

$$f(x, y) \neq f_1(x)\, f_2(y)$$

or equivalently, $F(x, y) \neq F_1(x)\, F_2(y)$.

2.34. Let X have the density function

$$f(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find a function $Y = h(X)$ which has the density function

$$g(y) = \begin{cases} 12y^3(1-y^2) & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

We assume that the unknown function h is such that the intervals $X \leqq x$ and $Y \leqq y = h(x)$ correspond in a one-one, continuous fashion. Then $P(X \leqq x) = P(Y \leqq y)$, i.e. the distribution functions of X and Y must be equal. Thus, for $0 < x, y < 1$,

$$\int_0^x 6u(1-u)\, du = \int_0^y 12v^3(1-v^2)\, dv$$

or

$$3x^2 - 2x^3 = 3y^4 - 2y^6$$

By inspection, $x = y^2$ or $y = h(x) = +\sqrt{x}$ is a solution, and this solution has the desired properties. Thus $Y = +\sqrt{X}$.

2.35. Find the density function of $U = XY$ if the joint density function of X and Y is $f(x, y)$.

Method 1.

Let $U = XY$ and $V = X$, corresponding to which $u = xy$, $v = x$ or $x = v$, $y = u/v$. Then the Jacobian is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ v^{-1} & -uv^{-2} \end{vmatrix} = -v^{-1}$$

Thus the joint density function of U and V is

$$g(u, v) = \frac{1}{|v|} f\left(v, \frac{u}{v}\right)$$

from which the marginal density function of U is obtained as

$$g(u) = \int_{-\infty}^{\infty} g(u, v)\, dv = \int_{-\infty}^{\infty} \frac{1}{|v|} f\left(v, \frac{u}{v}\right) dv$$

Method 2.

The distribution function of U is

$$G(u) = \iint_{xy \leqq u} f(x, y)\, dx\, dy$$

For $u \geqq 0$ the region of integration is shown shaded in Fig. 2-29. We see that

$$G(u) = \int_{-\infty}^{0} \left[\int_{u/x}^{\infty} f(x, y)\, dy \right] dx + \int_{0}^{\infty} \left[\int_{-\infty}^{u/x} f(x, y)\, dy \right] dx$$

Differentiating with respect to u using Leibniz's rule, we obtain

$$g(u) = \int_{-\infty}^{0} \left(\frac{-1}{x}\right) f\left(x, \frac{u}{x}\right) dx + \int_{0}^{\infty} \frac{1}{x} f\left(x, \frac{u}{x}\right) dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f\left(x, \frac{u}{x}\right) dx$$

The same result is obtained for $u < 0$, when the region of integration is bounded by the dashed hyperbola in Fig. 2-29.

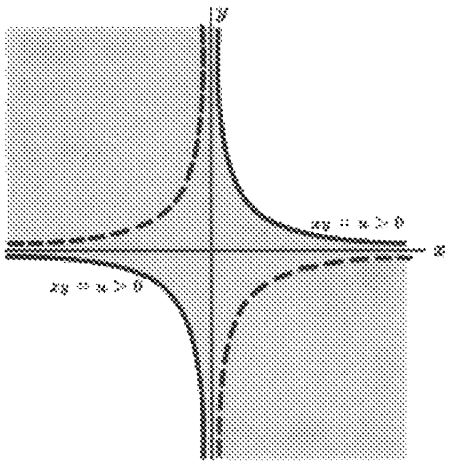


Fig. 2-29

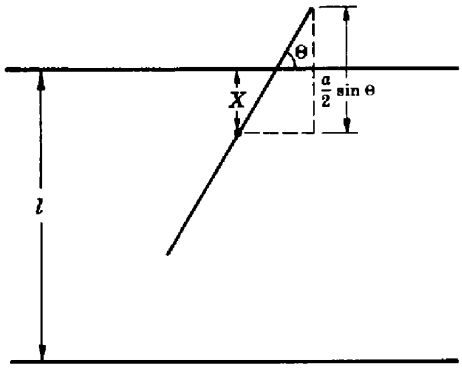


Fig. 2-30

2.36. A floor has parallel lines on it at equal distances l from each other. A needle of length $a < l$ is dropped at random onto the floor. Find the probability that the needle will intersect a line. (This problem is known as *Buffon's needle problem.*)

Let X be a random variable which gives the distance of the midpoint of the needle to the nearest line (Fig. 2-30). Let Θ be a random variable which gives the acute angle between the needle (or its extension) and the line. We denote by x and θ any particular values of X and Θ. It is seen that X can take on any value between 0 and $l/2$, so that $0 \leqq x \leqq l/2$. Also Θ can take on any value between 0 and $\pi/2$. It follows that

$$P(x < X \leqq x + dx) = \frac{2}{l}dx \qquad P(\theta < \Theta \leqq \theta + d\theta) = \frac{2}{\pi}d\theta$$

i.e. the density functions of X and Θ are given by $f_1(x) = 2/l$, $f_2(\theta) = 2/\pi$. As a check we note that

$$\int_0^{l/2} \frac{2}{l}dx = 1 \qquad \int_0^{\pi/2} \frac{2}{\pi}d\theta = 1$$

Since X and Θ are independent the joint density function is

$$f(x, \theta) = \frac{2}{l} \cdot \frac{2}{\pi} = \frac{4}{l\pi}$$

From Fig. 2-30 it is seen that the needle actually hits a line when $X \leqq (a/2)\sin\theta$. The probability of this event is given by

$$\frac{4}{l\pi}\int_{\theta=0}^{\pi/2}\int_{x=0}^{(a/2)\sin\theta} dx\, d\theta = \frac{2a}{l\pi}$$

When the above expression is equated to the frequency of hits observed in actual experiments, accurate values of π are obtained. This indicates that the probability model described above is appropriate.

2.37. Two people agree to meet between 2:00 P.M. and 3:00 P.M. with the understanding that each will wait no longer than 15 minutes for the other. What is the probability that they will meet?

Let X and Y be random variables representing the times of arrival, measured in fractions of an hour after 2:00 P.M., of the two people. Assuming that equal intervals of time have equal probabilities of arrival the density functions of X and Y are given respectively by

$$f_1(x) = \begin{cases} 1 & 0 \leqq x \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_2(y) = \begin{cases} 1 & 0 \leqq y \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

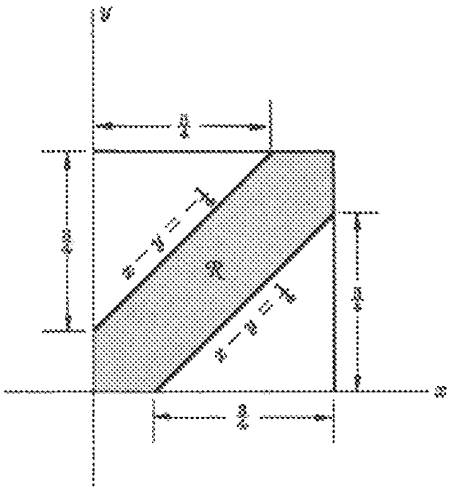


Fig. 2-31

Then, since X and Y are independent, the joint density function is

$$(1) \qquad f(x, y) = f_1(x)\, f_2(y) = \begin{cases} 1 & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

Since 15 minutes $= \frac{1}{4}$ hour, the required probability is

$$(2) \qquad P\left(|X - Y| \leqq \frac{1}{4}\right) = \iint_{\mathcal{R}} dx\, dy$$

where $\mathcal{R}$ is the region shown shaded in Fig. 2-31. The right side of (2) is the area of this region, which is equal to $1 - (\frac{3}{4})(\frac{3}{4}) = \frac{7}{16}$, since the square has area 1, while the two corner triangles have areas $\frac{1}{2}(\frac{3}{4})(\frac{3}{4})$ each. Thus the required probability is 7/16.

Supplementary Problems

DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

2.38. A coin is tossed three times. If X is a random variable giving the number of heads which arise, (*a*) construct a table showing the probability distribution of X and (*b*) graph the distribution.

2.39. An urn holds 5 white and 3 black marbles. If two marbles are drawn at random without replacement and X denotes the number of white marbles, (*a*) find the probability distribution for X and (*b*) graph the distribution.

2.40. Work Problem 2.39 if the marbles are drawn with replacement.

2.41. Let Z be a random variable giving the number of heads minus the number of tails in two tosses of a fair coin. Find the probability distribution of Z. Compare with the results of Examples 2.1 and 2.2.

2.42. Let X be a random variable giving the number of aces in a random draw of 4 cards from an ordinary deck of 52 cards. (*a*) Construct a table showing the probability distribution of X and (*b*) graph the distribution.

DISCRETE DISTRIBUTION FUNCTIONS

2.43. The probability function of a random variable X is shown in Table 2-7. (*a*) Construct a table giving the distribution function of X and (*b*) graph this distribution function.

Table 2-7

x	1	2	3
$f(x)$	1/2	1/3	1/6

Table 2-8

x	1	2	3	4
$F(x)$	1/8	3/8	3/4	1

2.44. Obtain the distribution function and its graph for (*a*) Problem 2.38, (*b*) Problem 2.39, (*c*) Problem 2.40.

2.45. Obtain the distribution function and its graph for (*a*) Problem 2.41, (*b*) Problem 2.42.

2.46. Table 2-8 shows the distribution function of a random variable X. Determine (*a*) the probability function, (*b*) $P(1 \leq X \leq 3)$, (*c*) $P(X \geq 2)$, (*d*) $P(X < 3)$, (*e*) $P(X > 1.4)$.

CONTINUOUS RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

2.47. A random variable X has density function

$$f(x) = \begin{cases} ce^{-3x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Find (*a*) the constant c, (*b*) $P(1 < X < 2)$, (*c*) $P(X \geq 3)$, (*d*) $P(X < 1)$.

2.48. Find the distribution function for the random variable of Problem 2.47. Graph the density and distribution functions, describing the relationship between them.

2.49. A random variable X has density function

$$f(x) = \begin{cases} cx^2 & 1 \leq x \leq 2 \\ cx & 2 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

Find (*a*) the constant c, (*b*) $P(X > 2)$, (*c*) $P(1/2 < X < 3/2)$.

2.50. Find the distribution function for the random variable X of Problem 2.49.

2.51. The distribution function of a random variable X is given by

$$F(x) = \begin{cases} cx^3 & 0 \leq x < 3 \\ 1 & x \geq 3 \\ 0 & x < 0 \end{cases}$$

If $P(X=3) = 0$ find (*a*) the constant c, (*b*) the density function, (*c*) $P(X > 1)$, (*d*) $P(1 < X < 2)$.

2.52. Can the function

$$F(x) = \begin{cases} c(1-x^2) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

be a distribution function? Explain.

2.53. Let X be a random variable having density function

$$f(x) = \begin{cases} cx & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (*a*) the value of the constant c, (*b*) $P(\frac{1}{2} < X < \frac{3}{2})$, (*c*) $P(X > 1)$, (*d*) the distribution function.

JOINT DISTRIBUTIONS AND INDEPENDENT VARIABLES

2.54. The joint probability function of two discrete random variables X and Y is given by $f(x, y) = cxy$ for $x = 1, 2, 3$ and $y = 1, 2, 3$, and equals zero otherwise. Find (*a*) the constant c, (*b*) $P(X = 2, Y = 3)$, (*c*) $P(1 \leq X \leq 2, Y \leq 2)$, (*d*) $P(X \geq 2)$, (*e*) $P(Y < 2)$, (*f*) $P(X = 1)$, (*g*) $P(Y = 3)$.

2.55. Find the marginal probability functions of (*a*) X and (*b*) Y for the random variables of Problem 2.54. (*c*) Determine whether X and Y are independent.

2.56. Prove that the discrete random variables X and Y are independent if and only if *for all values of* x *and* y their joint probability function can be expressed as the product of a function of x alone and a function of y alone.

2.57. Let X and Y be continuous random variables having joint density function

$$f(x, y) = \begin{cases} c(x^2 + y^2) & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Determine (*a*) the constant c, (*b*) $P(X < \frac{1}{2}, Y > \frac{1}{2})$, (*c*) $P(\frac{1}{4} < X < \frac{3}{4})$, (*d*) $P(Y < \frac{1}{2})$, (*e*) whether X and Y are independent.

2.58. Find the marginal distribution functions (*a*) of X and (*b*) of Y for the density function of Problem 2.57.

2.59. Prove the result of Problem 2.56 for the case where the variables are continuous.

CONDITIONAL DISTRIBUTIONS AND DENSITY FUNCTIONS

2.60. Find the conditional probability function (*a*) of X given Y, (*b*) of Y given X, for the distribution of Problem 2.54.

2.61. Let $$f(x, y) = \begin{cases} x + y & 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional density function of (*a*) X given Y, (*b*) Y given X.

2.62. Find the conditional density of (*a*) X given Y, (*b*) Y given X, for the distribution of Problem 2.57.

2.63. Let $$f(x, y) = \begin{cases} e^{-(x+y)} & x \geqq 0,\ y \geqq 0 \\ 0 & \text{otherwise} \end{cases}$$

be the joint density function of X and Y. Find the conditional density function of (*a*) X given Y, (*b*) Y given X.

CHANGE OF VARIABLES

2.64. Let X have density function

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leqq 0 \end{cases}$$

Find the density function of $Y = X^2$.

2.65. (*a*) If the density function of X is $f(x)$ find the density function of X^3. (*b*) Illustrate the result in part (*a*) by choosing

$$f(x) = \begin{cases} 2e^{-2x} & x \geqq 0 \\ 0 & x < 0 \end{cases}$$

and check the answer.

2.66. Prove Theorem 2-2, page 46.

2.67. Prove Theorem 2-3, page 46, for the cases (*a*) $\psi'(u) < 0$, (*b*) $\psi'(u) \geqq 0$, (*c*) $\psi'(u) \leqq 0$.

2.68. If X has density function $f(x) = (2\pi)^{-1/2} e^{-x^2/2}$, $-\infty < x < \infty$, find the density function of $Y = X^2$.

2.69. Verify that the integral of $g_1(u)$ in Method 1 of Problem 2.21 is equal to 1.

2.70. If the density of X is $f(x) = 1/\pi(x^2 + 1)$, $-\infty < x < \infty$, find the density of $Y = \tan^{-1} X$.

2.71. Complete the work needed to find $g_1(u)$ in Method 2 of Problem 2.21 and check your answer.

2.72. Let the density of X be

$$f(x) = \begin{cases} 1/2 & -1 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the density of (*a*) $3X - 2$, (*b*) $X^3 + 1$.

2.73. Check by direct integration the joint density function found in Problem 2.22.

2.74. Let X and Y have joint density function

$$f(x,y) = \begin{cases} e^{-(x+y)} & x \geq 0,\ y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

If $U = X/Y$, $V = X + Y$, find the joint density function of U and V.

2.75. Use Problem 2.22 to find the density function of (*a*) $U = XY^2$, (*b*) $V = X^2Y$.

2.76. Let X and Y be random variables having joint density function $f(x,y) = (2\pi)^{-1}e^{-(x^2+y^2)}$, $-\infty < x < \infty$, $-\infty < y < \infty$. If R and Θ are new random variables such that $X = R\cos\Theta$, $Y = R\sin\Theta$, show that the density function of R is

$$g(r) = \begin{cases} re^{-r^2/2} & r \geq 0 \\ 0 & r < 0 \end{cases}$$

2.77. Let

$$f(x,y) = \begin{cases} 1 & 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

be the joint density function of X and Y. Find the density function of $Z = XY$.

CONVOLUTIONS

2.78. Let X and Y be identically distributed independent random variables with density function

$$f(t) = \begin{cases} 1 & 0 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $X + Y$ and check your answer.

2.79. Let X and Y be identically distributed independent random variables with density function

$$f(t) = \begin{cases} e^{-t} & t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of $X + Y$ and check your answer.

2.80. Prove that (*a*) $f_1 * (f_2 + f_3) = f_1 * f_2 + f_1 * f_3$, (*b*) $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$.

2.81. Work Problem 2.21 by first making the transformation $2Y = Z$ and then using convolutions to find the density function of $U = X + Z$.

2.82. If the independent random variables X_1 and X_2 are identically distributed with density function

$$f(t) = \begin{cases} te^{-t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

find the density function of $X_1 + X_2$.

APPLICATIONS TO GEOMETRIC PROBABILITY

2.83. Two points are chosen at random on a line segment whose length is $a > 0$. Find the probability that the three line segments thus formed can be the sides of a triangle.

2.84. It is known that a bus will arrive at random at a certain location sometime between 3:00 P.M. and 3:30 P.M. A man decides that he will go at random to this location between these two times and will wait at most 5 minutes for the bus. If he misses it he will take the subway. What is the probability that he takes the subway?

2.85. Two line segments, AB and CD, have lengths 8 and 6 units respectively. Two points P and Q are chosen at random on AB and CD respectively. Show that the probability that the area of a triangle having height AP and base CQ is greater than 12 square units is equal to $(1 - \ln 2)/2$.

MISCELLANEOUS PROBLEMS

2.86. Suppose that $f(x) = c/3^x$, $x = 1, 2, \ldots$, is the probability function for a random variable X. (*a*) Determine c. (*b*) Find the distribution function. (*c*) Graph the probability function and the distribution function. (*d*) Find $P(2 \leq X < 5)$. (*e*) Find $P(X \geq 3)$.

2.87. Suppose that

$$f(x) = \begin{cases} cxe^{-2x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is the density function for a random variable X. (*a*) Determine c. (*b*) Find the distribution function. (*c*) Graph the density function and the distribution function. (*d*) Find $P(X \geq 1)$. (*e*) Find $P(2 \leq X < 3)$.

2.88. The probability function of a random variable X is given by

$$f(x) = \begin{cases} 2p & x = 1 \\ p & x = 2 \\ 4p & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

where p is a constant. Find (*a*) $P(0 \leq X < 3)$, (*b*) $P(X > 1)$.

2.89. (*a*) Prove that for a suitable constant c,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ c(1 - e^{-x})^2 & x > 0 \end{cases}$$

is the distribution function for a random variable X, and find this c. (*b*) Determine $P(1 < X < 2)$.

2.90. A random variable X has density function

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the density function of the random variable $Y = X^2$ and check your answer.

2.91. Two independent random variables, X and Y, have respective density functions

$$f(x) = \begin{cases} c_1 e^{-2x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad g(y) = \begin{cases} c_2 y e^{-3y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Find (*a*) c_1 and c_2, (*b*) $P(X + Y > 1)$, (*c*) $P(1 < X < 2, Y \geq 1)$, (*d*) $P(1 < X < 2)$, (*e*) $P(Y \geq 1)$.

2.92. In Problem 2.91 what is the relationship between the answers to (*c*), (*d*) and (*e*)? Justify your answer.

2.93. Let X and Y be random variables having joint density function

$$f(x, y) = \begin{cases} c(2x + y) & 0 < x < 1,\ 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find (*a*) the constant c, (*b*) $P(X > \frac{1}{2}, Y < \frac{3}{2})$, (*c*) the (marginal) density function of X, (*d*) the (marginal) density function of Y.

2.94. In Problem 2.93 is $P(X > \frac{1}{2}, Y < \frac{3}{2}) = P(X > \frac{1}{2})\,P(Y < \frac{3}{2})$. Why?

2.95. In Problem 2.91 find the density function (*a*) of X^2, (*b*) of $X + Y$.

2.96. Let X and Y have joint density function

$$f(x, y) = \begin{cases} 1/y & 0 < x < y,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Determine whether X and Y are independent. (*b*) Find $P(X > \frac{1}{2})$. (*c*) Find $P(X < \frac{1}{2}, Y > \frac{1}{3})$. (*d*) Find $P(X + Y > \frac{1}{2})$.

2.97. Generalize (*a*) Problem 2.78 and (*b*) Problem 2.79 to three or more variables.

2.98. Let X and Y be identically distributed independent random variables having density function $f(u) = (2\pi)^{-1/2}e^{-u^2/2}$, $-\infty < u < \infty$. Find the density function of $Z = X^2 + Y^2$.

2.99. The joint probability function for the random variables X and Y is given in Table 2-9. (*a*) Find the marginal probability functions of X and Y. (*b*) Find $P(1 \leqq X < 3, Y \geqq 1)$. (*c*) Determine whether X and Y are independent.

Table 2-9

X \ Y	0	1	2
0	1/18	1/9	1/6
1	1/9	1/18	1/9
2	1/6	1/6	1/18

2.100. Suppose that the joint probability function of random variables X and Y is given by

$$f(x, y) = \begin{cases} cxy & 0 \leqq x \leqq 2,\ 0 \leqq y \leqq x \\ 0 & \text{otherwise} \end{cases}$$

(*a*) Determine whether X and Y are independent. (*b*) Find $P(\frac{1}{2} < X < 1)$. (*c*) Find $P(Y \geqq 1)$. (*d*) Find $P(\frac{1}{2} < X < 1, Y \geqq 1)$.

2.101. Let X and Y be independent random variables each having density function

$$f(u) = \frac{\lambda^u e^{-\lambda}}{u!} \qquad u = 0, 1, 2, \ldots$$

where $\lambda > 0$. Prove that the density function of $X + Y$ is

$$g(u) = \frac{(2\lambda)^u e^{-2\lambda}}{u!} \qquad u = 0, 1, 2, \ldots$$

2.102. What is the analogue of the convolution theorem for discrete random variables? Illustrate by using Problem 2.101.

2.103. A stick of length L is broken into two parts. What is the probability that one part will have a length of more than double the other? State clearly what assumptions are made. Discuss whether you believe these assumptions are realistic and how you might improve them if they are not.

2.104. A floor is made up of squares of side l. A needle of length $a < l$ is tossed onto the floor. Prove that the probability of the needle intersecting at least one side is equal to $a(4l - a)/\pi l^2$.

2.105. For a needle of given length what should be the side of a square in Problem 2.104 so that the probability of intersection is a maximum? Explain your answer.

2.106. Let
$$f(x, y, z) = \begin{cases} 24xy^2z^3 & 0 < x < 1,\ 0 < y < 1,\ 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$
be the joint density function of three random variables X, Y and Z. Find (a) $P(X > \frac{1}{2}, Y < \frac{1}{2}, Z > \frac{1}{2})$, (b) $P(Z < X + Y)$.

2.107. A cylindrical stream of particles, of radius a, is directed toward a hemispherical target ABC with center at O as indicated in Fig. 2-32. Assume that the distribution of particles is given by
$$f(r) = \begin{cases} 1/a & 0 < r < a \\ 0 & \text{otherwise} \end{cases}$$
where r is the distance from the axis OB. Show that the distribution of particles along the target is given by
$$g(\theta) = \begin{cases} \cos\theta & 0 < \theta < \pi/2 \\ 0 & \text{otherwise} \end{cases}$$
where θ is the angle which line OP (from O to any point P on the target) makes with the axis.

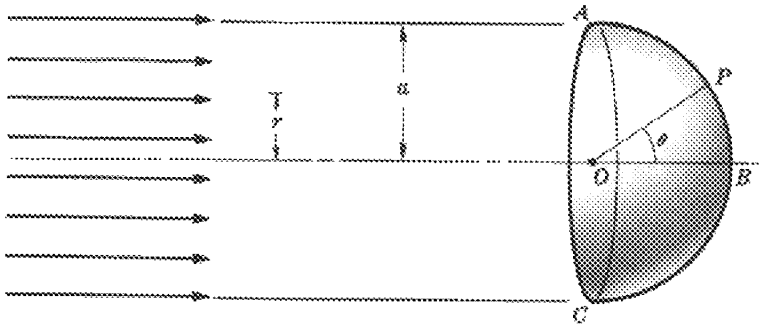


Fig. 2-32

2.108. In Problem 2.107 find the probability that a particle will hit the target between $\theta = 0$ and $\theta = \pi/4$.

2.109. Suppose that random variables X, Y and Z have joint density function
$$f(x, y, z) = \begin{cases} 1 - \cos\pi x \cos\pi y \cos\pi z & 0 < x < 1,\ 0 < y < 1,\ 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$
Show that although any two of these random variables are independent, i.e. their marginal density function factors, all three are not independent.

https://www.facebook.com/GrinchaZ